Computational crystal structure prediction and experimental characterisation of organic salts

Sharmarke Mohamed

A thesis submitted to UCL in partial fulfilment of the requirements for the degree of Doctor of Philosophy

June 2011

Department of Chemistry
University College London
20 Gordon Street
London
WC1H 0AJ
United Kingdom

I, Sharmarke Mohamed, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgements

Personal

My parents have provided more love and support than I can repay and this thesis is a tribute to their hard work. I am thankful for the healthy distraction my younger siblings have provided during the course of my graduate studies and I hope that my curiosity and appreciation of science has rubbed off on them. My older brother Abdi-Aziz has always put the interests of the family before his and I thank him for this. Robin Elmore has been a great teacher and has advised me wisely over the years. His interest in my work has given me the chance to communicate my research findings to a non-specialist and this has helped me clarify my thoughts during the drafting of this thesis. His willingness to proofread many parts of this thesis is greatly appreciated.

Academic

The past three years have given me the opportunity to work with many dedicated scientists and this has enriched my understanding of chemistry as well as the context in which scientific research is performed at university. Any list of these individuals would have to include my primary supervisor, Prof. Sally Price, at the top. She has been a constant source of guidance throughout. Because of her guidance, I've learned the importance of asking the right questions and I am more appreciative of the underlying assumptions behind the modelling methods used in this thesis. Prof. Derek Tocher has acted as my experimental supervisor and I have developed good experimental skills in solid form screening and single crystal X-ray crystallography under his leadership. Dr. Louise Price has provided invaluable administrative assistance with so many things and I am glad to have had her only a few feet from my desk at all times. Her organisation and commitment to her work has been a source of inspiration. The overlap between the research efforts of Mr. Nizar Issa and my own work on organic salts has provided plenty of opportunities for us to collaborate and bounce ideas off one another and this has been a useful experience. Dr. Gareth Welch was always willing to solve my computing problems, sometimes to the detriment of his own work and I thank him for his patience with me. Dr. Robert Lancaster has been a good source of advice from the very beginning. Dr. Doris Braun has provided constructive criticism of my ideas and writing, for which I am thankful. Dr. Panos Karamertzanis has had a direct contribution

to the research presented in this thesis and is acknowledged for performing the electronic structure calculations reported in Chapter 4 of this thesis. Dr. Abil Aliev is acknowledged for performing dynamic ^{1}H NMR experiments (Chapter 4). Dr. Martin Vickers is acknowledged for performing Rietveld refinement on the powder diffraction data of a mixture of solid forms reported in Chapter 4 of this thesis. Dr. Steve Firth is acknowledged for providing support with DSC/TG experiments. Prof. Alistair Florence and Dr. Jean-Baptiste Arlin are acknowledged for performing the experimental crystallisation screens on amantadine hydrochloride and memantine hydrochloride (Chapter 6). Finally, I would like to take this opportunity to thank the UCL Chemistry department for awarding me the DTA studentship that made this research possible.

Abstract

Approximately half of all pharmaceutical drugs are marketed as salts. This thesis pioneers the application of computational crystal structure prediction to organic salts containing the commonly used chloride or carboxylate counterions, and assesses the extent to which the theoretical calculations can be used to aid experimental efforts targeting organic salts.

A screen for multi-component solid forms of pyridine and 4-dimethylaminopyridine (DMAP) using a range of dicarboxylic acids led to one novel cocrystal of pyridine and six novel salts of DMAP. All novel crystal structures were solved using single crystal X-ray diffraction. At a simplistic level, salts differ from cocrystals in the position of the acidic proton within the crystal. For a selected set of structures, periodic *ab initio* calculations were shown to be useful in suggesting the observed N^{+}-H (salt) or O-H (cocrystal) covalent bond as the preferred atomic connectivity. Modelling the same crystal structures by lattice energy minimisation using a distributed multipole based electrostatic model proved successful if the correct proton connectivity was used. The observed structures of a model salt, cocrystal and disordered salt-cocrystal system were found to be the most stable or almost equi-energetic with the most stable structure in the predicted crystal energy landscapes. When the predictions were repeated using molecular structures with the wrong proton connectivity, the energetic ranking of the structure got worse.

The computational model of crystal structure prediction was successfully used to rationalise the different polymorphic and hydration behaviour of the pharmacologically active amantadine hydrochloride and memantine hydrochloride salts. Finally, a similar methodology was applied to 1,8-naphthyridinium fumarate and the calculations performed as part of the fifth international blind test of crystal structure prediction. Overall, the success in modelling the crystal structures of carboxylate and chloride salts illustrates the promise of crystal structure prediction in aiding experimental efforts of organic salt selection.

Summary of publications

Salt or Cocrystal?

1. Mohamed, S.; Tocher, D. A.; Vickers, M.; Karamertzanis, P. G.; Price, S. L. Salt or cocrystal? A new series of crystal structures formed from simple pyridines and carboxylic acids. *Cryst. Growth Des.* **2009,** *9* (6), 2881-2889.
2. Mohamed, S; Price, S. L.; Tocher, D. A. Computational prediction of salt and cocrystal structures - Does a proton position matter? *Int. J. Pharm.* **2011**, doi: 10.1016/j.ijpharm.2011.03.063.

Crystal structure prediction of chloride salts

3. Arlin, J. B; Florence, A. J.; Mohamed, S; Tocher, D. A.; Price, S. L. Using crystal structure prediction to rationalise the polymorphic and hydration behaviour of amantadine hydrochloride and memantine hydrochloride. *Cryst. Growth Des.* **2011**, *in preparation.*

Fifth international blind test of crystal structure prediction

4. Bardwell, D; Mohamed, S *et al*. Fifth international blind test in crystal structure prediction. *Acta Crystallogr., Sect. B.* **2011**, *in preparation.*

Polymorphism

5. Mohamed, S.; Barnett, S. A.; Tocher, D. A.; Price, S. L.; Shankland, K.; Leech, C. K. Discovery of three polymorphs of 7-fluoroisatin reveals challenges in using computational crystal structure prediction as a complement to experimental screening. *CrystEngComm* **2008,** *10* (4), 399-404.
6. Shankland, K.; Leech, C. K.; Mohamed, S.; Barnett, S. A.; Tocher, D. A. 7-Fluoroisatin-1,4-dioxane(1/1). *Acta Crystallogr. , Sect. E* **2007,** *63*, O3574-U4110.
7. Mohamed, S.; Barnett, S. A.; Tocher, D. A. 7-Fluoroisatin-dimethyl sulfoxide (1/1). *Acta Crystallogr. , Sect. E* **2007,** *63*, O3575-U4120.
8. Mohamed, S.; Barnett, S. A.; Tocher, D. A. 5-Fluoroisatin-dimethyl sulfoxide (1/1). *Acta Crystallogr. , Sect. E* **2007,** *63*, O3576-U4130.

9. Mohamed, S.; Barnett, S. A.; Tocher, D. A. 5-fluoro-3-hydroxy-3-(nitromethyl)-1H-indol-2(3H)-one. *Acta Crystallogr. , Sect. E* **2007,** *63*, O3577-U4140.

Table of contents

Figures

Tables

Schemes

1 Introduction

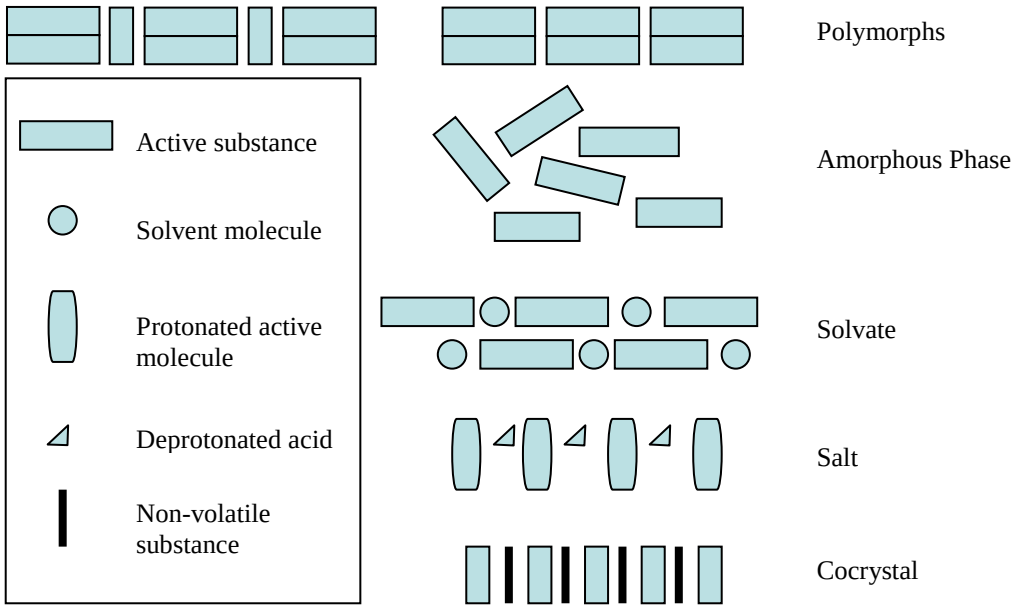
1.1 Overview

This chapter provides a brief introduction to the possible solid forms that can result from crystallisation of a molecule with one or more ionisable functional groups. The emphasis will be on organic salts, which are a subset of a larger family of solid forms known as multi-components. An overview of the experimental crystallisation methods used in this thesis will be provided and the basic principles of crystal engineering used throughout this thesis will be introduced. The industrial salt selection process will be outlined and the regulatory requirements for patenting salts of existing drugs will be discussed. An introduction to computational crystal structure prediction will be given but the discussion will be limited to the insights that can be gained from computational modelling when used alongside experimental screening. This is pertinent because much of this thesis (Chapters 4-6) will use the results of salt crystallisation screens to validate a computational model of crystal structure prediction based on the global minimisation of the lattice energy. A more detailed account of the theoretical basis of crystal structure prediction is provided in Chapter 2 of the thesis.

1.2 Solid form diversity

1.2.1 Introduction

When formulating the optimal solid form of an active pharmaceutical ingredient (API), pharmaceutical companies consider a variety of possible crystal forms. Often this is motivated by the fact that the API has undesirable[1,2] physiochemical properties and the aim is to perform a sufficiently extensive set of crystallisation screens[3] to find a crystal form that has the desired physicochemical profile. Scheme 1.1 shows the possible solid forms that may result from the screen. We assume that the active pharmaceutical ingredient contains one or more ionisable functional groups (assumed to be basic in Scheme 1.1). Polymorphic and amorphous phases of the API are considered from the outset. If it is not possible to crystallise the API as a polymorph or amorphous phase, or the physicochemical profile of the API is not improved with any polymorph or amorphous phase found, multi-component solid forms of the drug are crystallised. By



Scheme 1.1: Schematic depiction of the possible solid state forms that may be adopted by a given active substance. Reproduced from *Polymorphism in the Pharmaceutical Industry*[4].

definition multi-component crystal forms are those that contain the API and at least one other chemically distinguishable species in the crystal structure. Examples of multi-component solid forms (Scheme 1.1) are *solvates*, *salts* and *cocrystals*. Each of these multi-component solid forms is capable of displaying polymorphism and so the variety of possible crystal forms can be larger than depicted in Scheme 1.1. Cocrystals have received a great deal of attention[5-7] in recent years. It is possible[8] to maximise the *likelihood* of forming cocrystals from two reagents by considering the most important intermolecular interactions when the two components crystallise. This is in the context of the emerging field of crystal engineering[9] which uses hydrogen bond complementarity (see section 1.2.2) between a pair of molecules to decide if they are likely to cocrystallise. A more detailed discussion of the importance of hydrogen bonding in driving molecular complementarity as well as the limitations of the so called *supramolecular synthon approach*[10,11] can be found in section 3.2 of this thesis. Unfortunately, there has been considerable debate[12,13] in the literature surrounding the definition of a cocrystal. For example, Aakeroy[14] has suggested a cocrystal be defined as a multi-component crystal composed of neutral molecules that are solids at ambient conditions. This clearly has its merits since it distinguishes a cocrystal from other multi-component crystals where one of the molecular components is charged (salts) or is a solvent used in the crystallisation experiment (solvate). An alternative definition

proposed by Bond[15] suggests that the term *cocrystal* be used only as a synonym for a multi-component crystal composed of neutral molecules. Bond[15] argues that any attempt to make *cocrystal* more specific would only lead to further definitions since neither *solvate* nor Aakeroy's definition of a *cocrystal*, adequately describe crystals formed from two liquid reagents[16]. To avoid any possible confusion over terminology, the following definitions of *salt* and *cocrystal* will be used throughout this thesis.

- *Salt.* A salt is a multi-component crystal form whose asymmetric unit consists of a stoichiometric ratio of a negatively charged anion and a positively charged cation. The crystal may contain additional molecules (*i.e.* solvent or acid) that have been incorporated into the lattice during crystallisation.

- *Cocrystal.* A cocrystal is a structurally homogeneous crystalline material that contains two or more neutral building blocks that are structurally distinct from one another. The cocrystal formers[17] need not be solids[14] at ambient conditions.

1.2.2 Hydrogen bonds and graph sets

Hydrogen bonds[18] are important intermolecular interactions as they direct the relative orientation and hence packing of molecules/ions in organic crystals. When comparing two or more crystal structures, solid state chemists routinely consider the hydrogen bonds present as a means of distinguishing between them. The term *hydrogen bond* first appeared in the chemical literature in 1935 when Pauling used it to account for the residual entropy of ice[19]. He later defined[20] a hydrogen bond as an 'interaction that directs the association of a covalently bound hydrogen atom with one or more other atoms' and made it clear that its origins are electrostatic in nature. For a general configuration X-H···A, the hydrogen bond is illustrated by the dotted line while the solid line represents a covalent bond between X and H. The symbols X and A represent *hydrogen bond donor* and *hydrogen bond acceptor* atoms respectively. In the crystal structures of organic salts, the donor and acceptor atoms carry formal charges and the hydrogen bond is written as X^{+}-H···A^{-}.

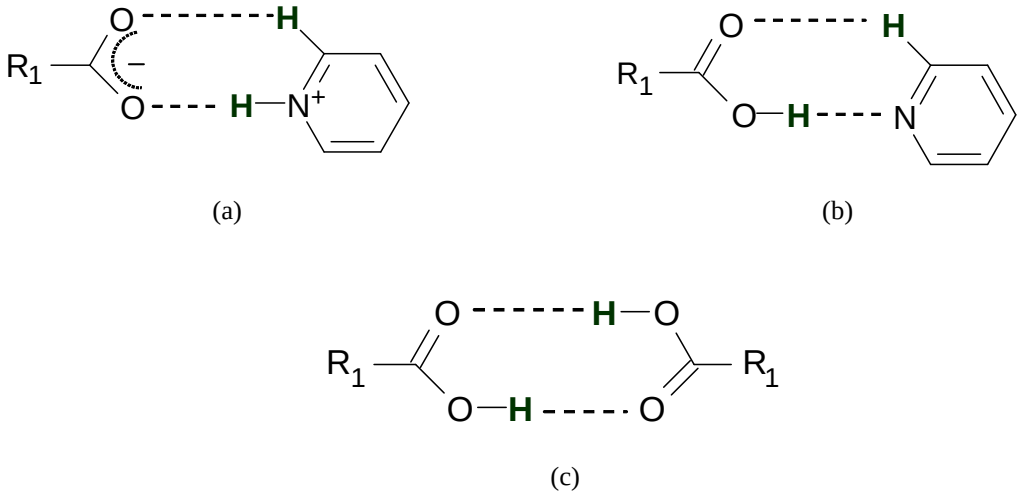
Ionic		
N^+-H···O^-	N^+-H···Cl^-	
Very strong		
F-H···F	P-OH···O=P	
Strong		
O-H···O=C	N-H···O=C	O-H···O-H
Weak		
C-H···O	C-H···N	C-H···F-C

Table 1.1: Examples of hydrogen bond interactions that fall under the categories ionic, very strong, strong or weak. Reproduced from *The Weak Hydrogen Bond*[21].

It is possible to classify[21] hydrogen bonds into one of four categories depending on the strength of the intermolecular interaction. The first category is *ionic* and occurs in the crystal structures of organic salts where we find that the molecules containing atoms X and A carry formal charges. The remaining three categories[21] are *very strong*, *strong* or *weak* and occur in structures where only one of the molecules that contain atoms X or A carries a formal charge or none of the molecules carry a formal charge. Examples of hydrogen bond interactions that fall under each of these categories are given in Table 1.1. Hydrogen bonds play an important role in the work of crystal engineers who are interested in controlling the structures and hence properties of organic crystals. Etter proposed[22] a set of rules for predicting the likely hydrogen bond interactions expected when a molecule crystallises and these can be summarised as follows:

1. All good proton donors and acceptors are used in hydrogen bonding
2. Six-membered-ring intramolecular hydrogen bonds form in preference to intermolecular hydrogen bonds
3. The best proton donors and acceptors remaining after intramolecular hydrogen-bond formation, form intermolecular hydrogen bonds to one another

Whilst these rules were originally developed for the crystal structures of organic molecules, they have since been extended to the design of cocrystal solid forms[23]. Crystal engineers usually talk about *supramolecular synthons*[24] or *hydrogen bond motifs*[25] when discussing the key structure directing interactions that occur in crystal structures. Examples of three common[26,27] hydrogen bond motifs found in organic



Scheme 1.2: Illustration of common hydrogen bond motifs found in pyridinium carboxylate salts (a), pyridine carboxylic acid cocrystals (b) and carboxylic acids (c). The hydrogen bonds are illustrated by the dotted lines and the hydrogen atoms involved in the motif are highlighted green.

crystal structures are illustrated in Scheme 1.2. The pyridinium carboxylate and pyridine carboxylic acid motifs shown in Scheme 1.2 can be qualitatively distinguished on the basis of the acidic proton position. However, assignment based on visual inspection of acidic proton connectivity is demanding of the quality of the X-ray diffraction data. An alternative[28] method of distinguishing between the salt and cocrystal hydrogen bond motifs shown in Scheme 1.2 is to look at the C-N-C angle and C-O bondlengths of the acid-base/cation-anion pair in the crystal. Zaworotko's survey[28] of the Cambridge Structural Database (CSD[29]) has shown an average C-N-C angle of 116(2)° for neutral pyridines. By contrast, when the pyridine is protonated at N, the average C-N-C angle increases to 121(2) °. The same work has shown that the two C-O bondlengths of a carboxylate group are comparable (due to electron delocalisation) and have a CSD average of 1.25(2) Å. By contrast a carboxylic acid group can be identified if it has one long C-O bond and one short C=O bond with CSD averages of 1.31(2) Å and 1.21(2) Å respectively. Chapter 4 of this thesis will investigate the effects of acidic proton in the crystal structures of salts and cocrystals using the results of computational modelling and experimental screening.

As part of her work on hydrogen bonds, Etter proposed a system[30] for describing hydrogen bonds in crystal structures. This system is based on *graph set* theory. A given hydrogen bond motif may be described by a graph set according to the general symbol

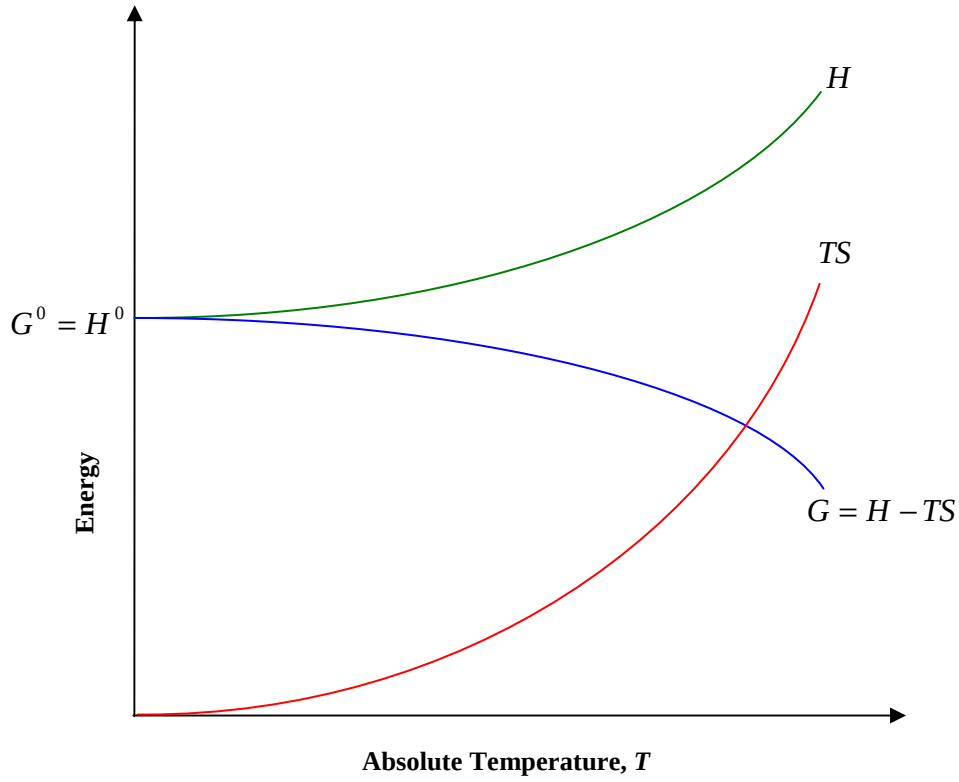
$$G_d^a(n)$$

where *G* is the descriptor of the hydrogen bond motif, *d* and *a* are the number of donor and acceptor atoms respectively and *n* is the degree of the pattern and describes the number of atoms in the repeat unit of the motif. By way of example, the carboxylic acid dimer motif shown in Scheme 1.2 has the graph set $R_2^2(8)$. The descriptor *R* is taken to mean the motif consists of a *ring*. There are 2 (O-H) donors and 2 (C=O) acceptors in the dimer, hence the numerical values of the subscript and superscript. The repeat unit is 8 because there are 8 atoms in the dimer. Using the same principles, we can assign an $R_2^2(7)$ graph set to the pyridine carboxylic and pyridinium carboxylate motifs shown in Scheme 1.2.

1.2.3 Polymorphism

Polymorphism[31], as applied to the crystal structures of organic compounds, may be defined as the ability of a substance to crystallise in more than one solid form. The term derives its meaning from the Greek (*poly*=many and *morph*=forms), and finds applicability in the life sciences where one finds terms like 'genetic polymorphism' to mean the presence of at least two phenotypes in a population, due to the expression of different alleles of a given gene. In crystallography, it has a more rigorous definition[31] and is only valid when a compound displays at least two distinct crystal forms. Crystal polymorphism may be caused by differences in the packing or conformations adopted by molecules or ions. The phenomenon of polymorphism can be traced as far back as the late 18th century when Klaproth realised that the minerals aragonite and calcite displayed the same chemical composition, $CaCO_3$, and further work by Mitscherlich[32] on this system as well as other minerals, verified conclusively the existence in nature of different forms of the same species. When polymorphism is discovered late in the drug development process, this can be hugely costly to pharmaceutical companies[33]. Much of the early developments in computational crystal structure prediction[34,35] were motivated by the challenges posed by polymorphism.

The molecules/ions in the crystal lattice of polymorphs differ in the nature of the



Scheme 1.3: Energy-Temperature diagram of a solid under constant pressure.

non-covalent interactions, such as hydrogen bonds, van der Waals forces and π-π stacking interactions. This leads to differences in the enthalpy of polymorphs. A systematic study of the sort of energy-temperature diagrams depicted in Scheme 1.3 - for the Helmholtz, *A*, and internal energy, *U*, surfaces - has led Burger *et al*[36] to conclude that the enthalpy of a crystalline solid under ambient pressure conditions has a negligible contribution from *pV* work. As a consequence we may approximate the enthalpy (*H*) according to

$$H = U + pV \approx U \qquad \textbf{(1)}$$

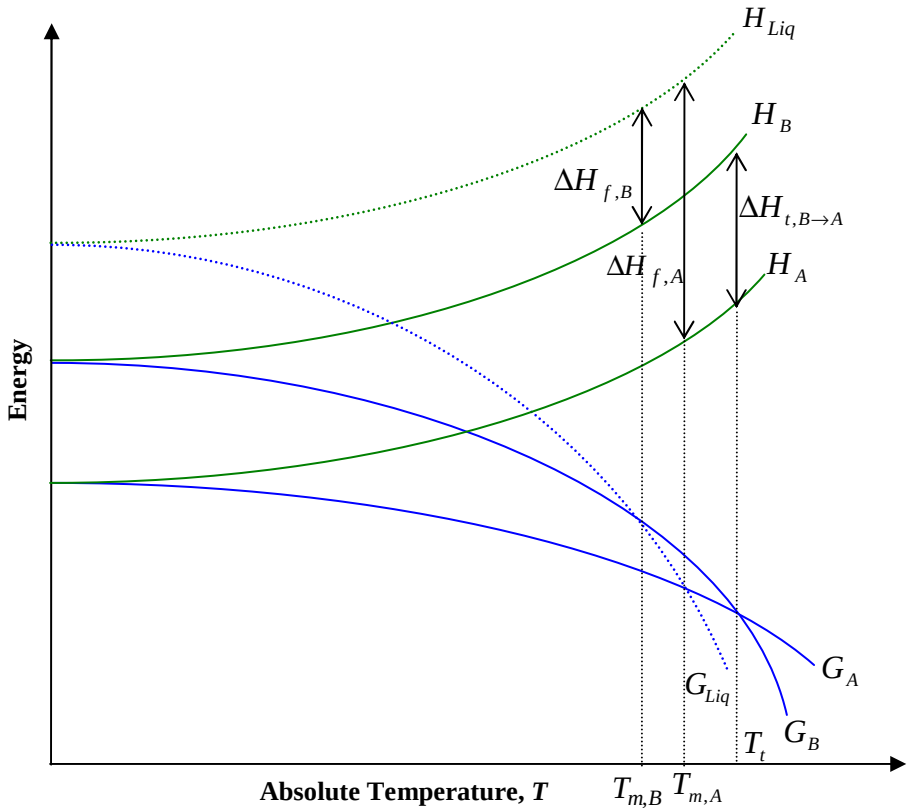
This approximation is relevant to the modelling in this thesis because it allows us to relate the crystal enthalpy (*H*) to the lattice energy (E_{latt}), through the assumption that the differences in the internal energy (*U*) are equal to the differences in the lattice energy (E_{latt}). This requires that the heat capacities of the structures being compared is equal, which is a reasonable assumption to make. The lattice energy refers to the energy of the static crystal at 0 K relative to the energies of the infinitely separated

molecules/ions. The lattice energy can be estimated (see section 2.5) as a sum of a repulsion-dispersion contribution and an electrostatic contribution obtained from a distributed multipole analysis of the *ab initio* wavefunction of the isolated ions. A detailed discussion of the various contributions towards the crystal lattice energy and the approximations made in modelling these can be found in Chapter 2 of the thesis.

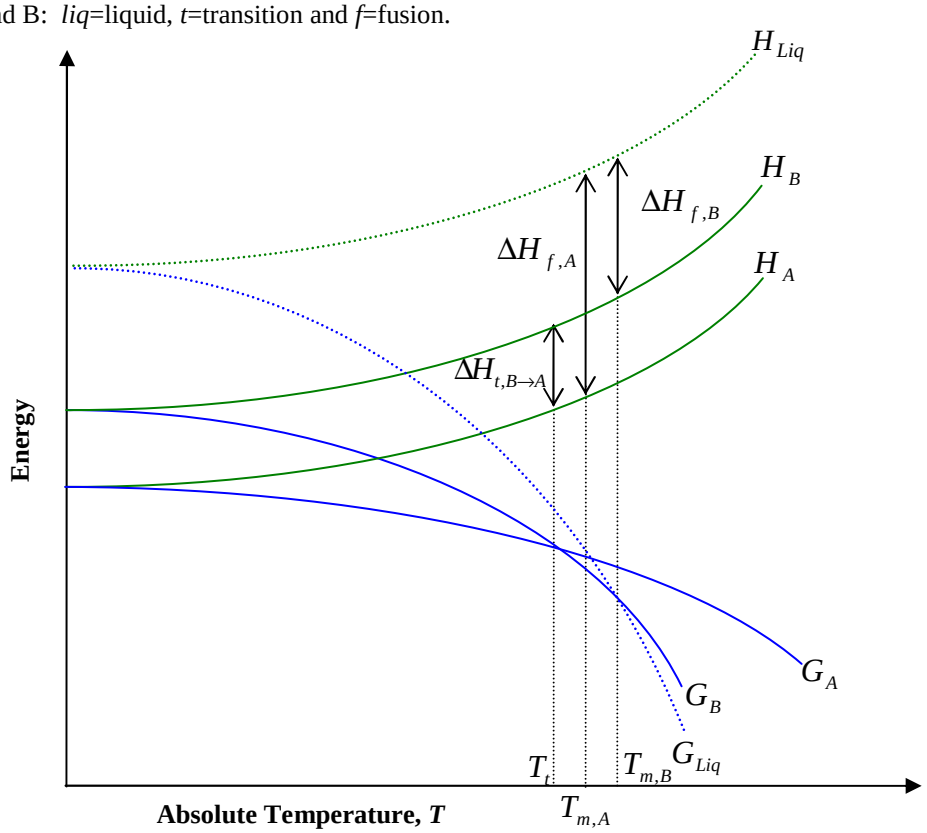
An idealised two-dimensional plot of the Gibbs and enthalpy surfaces as a function of temperature is given in Scheme 1.3. The entropic contribution towards the Gibbs free energy, *TS*, is shown as an additional line plot. The enthalpy isobar increases with temperature because the derivative of the curve is the isochoric heat capacity of the solid, which increases with temperature according to Einstein's formulation of the heat capacity for a monatomic crystal[37]. The entropic contribution, *TS*, is shown as increasing with temperature because a higher temperature gives atoms and molecules greater mobility, thereby increasing the entropy. The free energy falls with increasing temperature because the slope of the curve is the negative of the entropy. In this thesis, the lattice energies are true at 0 K and we neglect the zero-point contributions at this temperature. Thus all the crystal lattice energies reported in this thesis refer to calculations performed at the origin of the enthalpy, *H*, curve shown in Scheme 1.3.

The usefulness of Scheme 1.3 is that it may be used to rationalise the thermodynamic relationship between polymorphs. If we consider a system with two polymorphs, *A* and *B*, where form *A* is more stable than form *B* at absolute zero, two possible energy-temperature diagrams may be drawn. The first is for a monotropically related pair of polymorphs (Scheme 1.4). A pair of polymorphs are said to be monotropically related if there is no change in the stability ordering below the melting points of the two forms. In theory, metastable form *B* should be able to undergo a spontaneous exothermic transformation to the more stable form *A*, since this transformation is thermodynamically feasible at all temperatures below T_t. The larger activation barrier associated with solid-solid transformations means that in a monotropic system, such a transformation is not always practical.

A pair of polymorphs is said to be enantiotropically related (Scheme 1.5) if there is a transition point, T_t, below the melting point of both polymorphs, where the stability order changes. At the transition temperature, the Gibbs free energy of both polymorphs



Scheme 1.4: Schematic[4] energy-temperature diagram for a monotropically related pair of polymorphs, A and B: *liq*=liquid, *t*=transition and *f*=fusion.



Scheme 1.5: Schematic[4] energy-temperature diagram for an enantiotropically related pair of polymorphs. (*G* refers to the Gibbs free energy and *H* is the enthalpy).

is the same, and so the polymorphs are said to be in equilibrium with each other. For an enantiotropically related pair of polymorphs, Scheme 1.5 shows that below T_t, form B can undergo a spontaneous exothermic transformation into form A in agreement with the Gibbs free energy ordering of the two polymorphs.

1.3 Crystallisation

1.3.1 Factors controlling crystallisation

Generally speaking, the crystallisation of a pharmaceutically active substance has two simultaneous objectives: to purify the product so that trace compounds resulting from the synthesis can be removed and to obtain crystals of the pharmaceutically active substance. The crystallisation process is influenced by several factors. These include *supersaturation*, *seeding*, *cooling rate* and *attrition*.

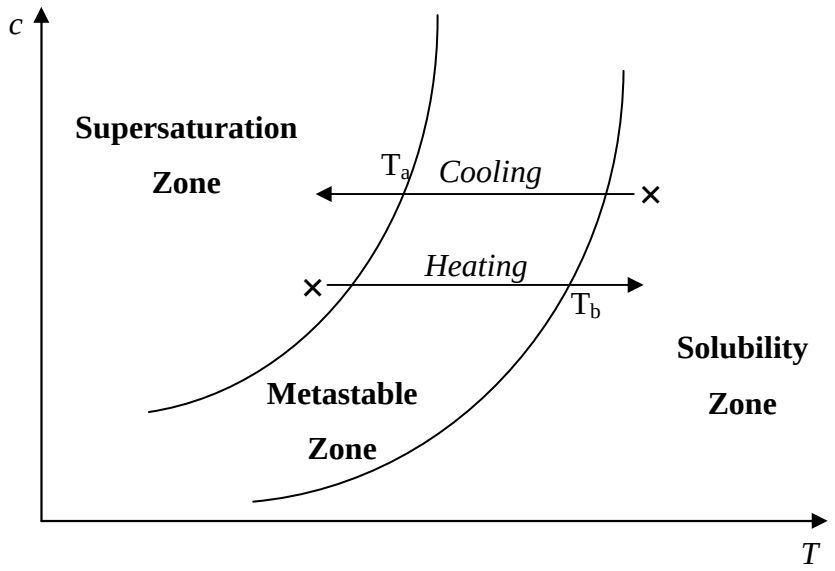
The driving force behind the crystallisation process is the level of *supersaturation*. When a solution is supersaturated it has excess solute for a given volume of solvent. Thus precipitation to form crystals can be achieved at a faster rate. The level of supersaturation can be quantified by calculating the difference in the concentration, *c*, of the solution and the concentration at equilibrium, *c**, at the same temperature

$$\Delta c = c - c^* \qquad \textbf{(2)}$$

This leads to the following formal definition for the level of supersaturation as the difference in the chemical potentials of the solution, μ, and the chemical potential of the solution at equilibrium, μ^*

$$\Delta\mu = \mu - \mu^* = kT\ln(c/c^*) \qquad \textbf{(3)}$$

where k is the Boltzmann constant and T is the absolute temperature of the experiment. Ostwald[38] first introduced the terms *labile* and *metastable* supersaturation to classify supersaturated solutions in which primary nucleation could and could not occur. Subsequent work by Miers and Isaac[39] led to a diagrammatic representation of



Scheme 1.6: Definition of the metastable zone in crystallisation experiments from solution.

Ostwald's thoughts (Scheme 1.6). When supersaturation increases on cooling, crystals spontaneously start to appear at a certain temperature, T_a. This is characteristic of primary nucleation. By changing the concentration several times and repeating the operation, the supersaturation curve is obtained (Scheme 1.6). By increasing the temperature from this supersaturated region, the crystals are dissolved at a temperature T_b and the original solution is obtained. This is repeated several times at different levels of supersaturation to obtain the solubility curve. The two curves in Scheme 1.6 define three regions[40]

1. The stable/solubility zone, where crystallisation cannot occur.
2. The metastable zone, between the solubility and supersaturation curves, where crystallisation may or may not occur.
3. The supersaturation zone at low temperatures, where there is excess solid and crystallisation is probable.

In general, the solubility curve is a thermodynamic limit, which depends only on the solute and solvent used, whereas the supersaturation curve is a kinetic limit which depends on the rate of cooling, stirring and the presence of impurities.

The effects of *seeding* affect the pathway as well as the quality of the resulting crystals. When seeding occurs near the solubility curve, it proceeds at low supersaturation. As a consequence there is little secondary nucleation, and large

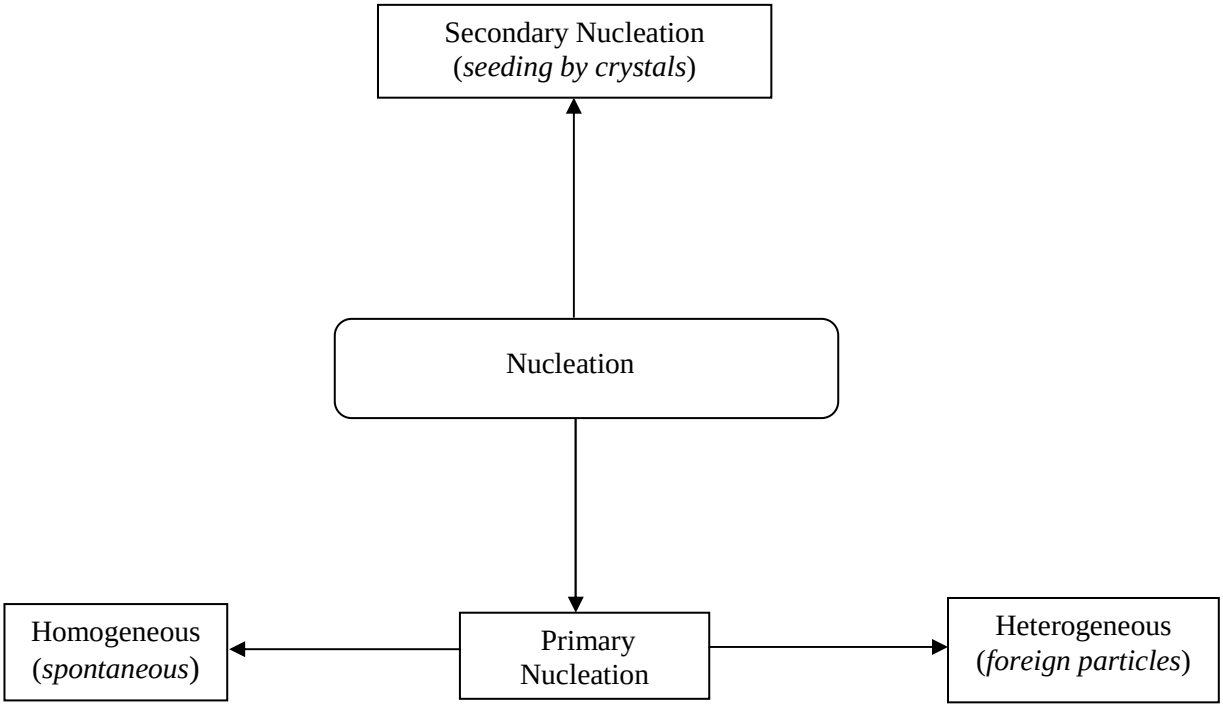
crystals are grown. By contrast, seeding near the supersaturation curve, leads to many secondary nucleation effects, and the resulting crystals are often small because of the rapid crystal growth rate. When a substance is known to display polymorphism, intentional[41] seeding of one polymorph with a batch containing the growing crystallites, is an efficient way of preferentially crystallising the seeded polymorph. Conversely, the presence of impurities can lead to unintentional seeding of the growing crystallites and this has been used[42] to explain cases of 'disappearing'[41,43] polymorphs.

The *cooling rate* is a key experimental parameter in the crystallisation process. Fast cooling will lead to the rapid crossing of the metastable zone and when the system reaches the supersaturation curve, the crystals that result from the nucleation will be small in size as a result of the many nuclei that have crashed out of solution before they can assemble with other nuclei to form large crystals. The opposite is true for slow cooling experiments and this is often the preferred mode of crystallisation at low temperatures.

Attrition arises as a consequence of particle collisions. It is proportional to the stirring rate and is a crystallisation parameter extensively employed by chemical engineers. The idea is that a rapid rate of stirring will favour molecular collisions in solution and hence lead to an increase in attrition. Stirring can also affect growth and nucleation kinetics as a consequence of the increase in seed proliferation.

1.3.2 Solution crystallisation methods

Crystallisation has for a long time played a central role in synthesis for the purification of substances. Organic chemists for example routinely perform the *re*-crystallisation of a substance once the crude product has been isolated, because the impurities that remain after dissolving the crude material can easily be filtered from the solution of interest. Crystallising this solution, often by cooling will then lead to a purified substance. With the development of methods like chromatography and chiral stationary phases, for the purification of substances, it has however lost its importance over the years, to the point where it is now mentioned in newly published chemistry textbooks only in passing[44]. Although many compounds emerge from preparative procedures in 'crystalline form', these crystals are often not suitable enough for structure determination via single crystal X-ray diffraction experiments. In these circumstances, the following well established



Scheme 1.7: Possible pathways by which nucleation can be achieved. Secondary nucleation is induced by crystals in the form of seeding. Homogeneous primary nucleation is spontaneous whereas heterogeneous primary nucleation is induced by foreign particles such as dust but in the absence of intentional seeding.

methods[45] can be used to prepare good quality single crystals.

- *Simple solvent evaporation*. The simplest method of crystallising a salt from a solution saturated with a stoichiometric ratio of acid and base is to allow the excess solvent to evaporate. An obvious limitation of this method is that it requires both the acid and base to be soluble in the solvent chosen otherwise there is the risk of the acid or base crystallising out of solution without forming a salt. Both the solvent and the rate of solvent evaporation can be varied and each parameter can give rise to crystals of different morphology that may or may not be polymorphs. A variant on simple solvent evaporation at room temperature is to carry out a slow cooling experiment. This sort of set up is particularly good for systems which display moderate solubility in the solvent of interest. In this case there is insufficient crystallising matter in solution and cooling will aid the crossing of the metastable zone and into the supersaturation zone (Scheme 1.6). In general, solvent evaporation leads to a crystalline phase because it increases the solute/solvent ratio of the saturated solution

and leads to a supersaturated solution. This supersaturated solution will lead to primary nucleation sites, which aid the production of further crystalline forms via precipitation.

- *Vapour diffusion*. The idea here is that the precipitant (known as the *anti-solvent*) will diffuse into a saturated solution containing the system of interest (in a closed set-up). Such a process will decrease the solubility of the salt in the saturated solution, thus initiating crystallisation. This is only possible if the anti-solvent is more volatile than the saturating solvent. Both the acid and base must be insoluble in the anti-solvent used and the two solvents must be miscible. The advantage of this method is that several tubes may be set up inside the same container and that very small quantities may be used. Often the acid and base have different solubilities in the solvent and anti-solvent, thereby limiting the applicability of this method to the crystallisation of salts.

- *Use of seed crystals*. If any of the above methods, produce crystals of reasonable quality but of too small a size, then these may be used as seed crystals for simple solvent evaporation or slow cooling experiments. The effect is often not deliberate[12]. However, seeding can be very effective when seed crystals are introduced within the metastable zone. The seed crystals affect the nucleation process and this sort of method falls under the secondary nucleation pathway described in Scheme 1.7. The seed crystals are transferred in the presence of some mother liquor as they lose their effectiveness when dry. As a general rule, the fewer the seed crystals, the larger the resulting crystals.

In addition to the above solution based methods for obtaining good quality single crystals of salts, mechanochemical grinding experiments can also be used to induce crystallisation and further details are given below.

1.3.3 Mechanochemical methods of crystallisation

Mechanochemical[46] methods of crystallisation are those that use mechanical grinding to initiate crystallisation between two molecular components. This method is particularly useful when solution based methods fail, as can happen when the acid and base have very different solubilities in the solution medium used. Over the past few years, mechanical grinding has proven to be an efficient[47,48] method in the supramolecular synthesis of cocrystal solid forms and there is evidence[49] suggesting it can be used to synthesise solid forms[50,51] that cannot be accessed via traditional solution crystallisation methods. A further contrast to the solution crystallisation methods described above is the fact that mechanochemical grinding is environmentally friendly[52] since it requires little or no solvent. Whilst it is true to say that in recent years grinding of acid-base pairs to produce salts has been much less explored than grinding to produce molecular cocrystals, the method is just as amenable to the synthesis of salt[53] solid forms.

Two types of mechanochemical grinding are routinely used within the crystal engineering community: *neat grinding* and *solvent drop grinding*. With neat grinding, the components are mixed and then ground. The grinding can be performed manually using a pestle and mortar or mechanically using a ball mill. In the case of solvent drop grinding, a small catalytic amount of solvent is added to the mixture and this is followed by grinding. The catalytic solvent offers[54] a crystallisation pathway with higher product yield and higher crystallinity than is the case with neat grinding. A recent review[54] into the mechanism of cocrystal formation via neat grinding has highlighted the importance of molecular diffusion in producing the final crystalline product. Whilst there is some reason[54] to speculate that solvent drop grinding also operates via a mechanism of molecular diffusion, it is not clear how the solvent contributes to the enhanced product yield observed with this method when compared to neat grinding. One of the negative aspects associated with mechanochemical grinding is the difficulty with product characterisation. The advantage of using solution crystallisation methods is that if successful, they lead to single crystals that are easy to characterise via X-ray diffraction. With mechanochemical grinding, single crystals of diffraction quality are not produced and the usual method of characterisation is X-ray powder diffraction. The recent work of Cruz-Cabeza *et al*[55] also shows that computational crystal structure prediction may be used as an aid to solving the crystal structures of powders obtained from grinding.

1.3.4 Novel crystallisation and high throughput methods

In addition to the above methods of crystallisation, which will be used throughout this thesis, there are other methods that have gained some traction among crystal engineers and industrial chemists. Most of the methods discussed below were developed for the screening of single component systems but are mentioned here for completeness. The *membrane crystallisation*[56] technique is an example of a novel method that has appeared in recent years. The method is so called because it uses a microporous membrane, which acts as a physical support that separates two isothermal solutions, and allows the solvent to meet between them. This method has the advantage of being able to control the rate of solvent evaporation. This in turn affects the morphology as well as crystallinity of the resulting products[57]. Di Profio has shown[58] that the polymorphs of glycine can be selectively crystallised using this technique. Ultrasonic irradiation[59] has been used as an external parameter that can affect the crystallisation process. The method has often been employed during the nucleation stage of the crystallisation process, and despite a lack of a detailed understanding of the mechanism behind the effect, it is commonly accepted that the ultrasonic irradiation allows the crystallisation to take place at lower supersaturation levels.

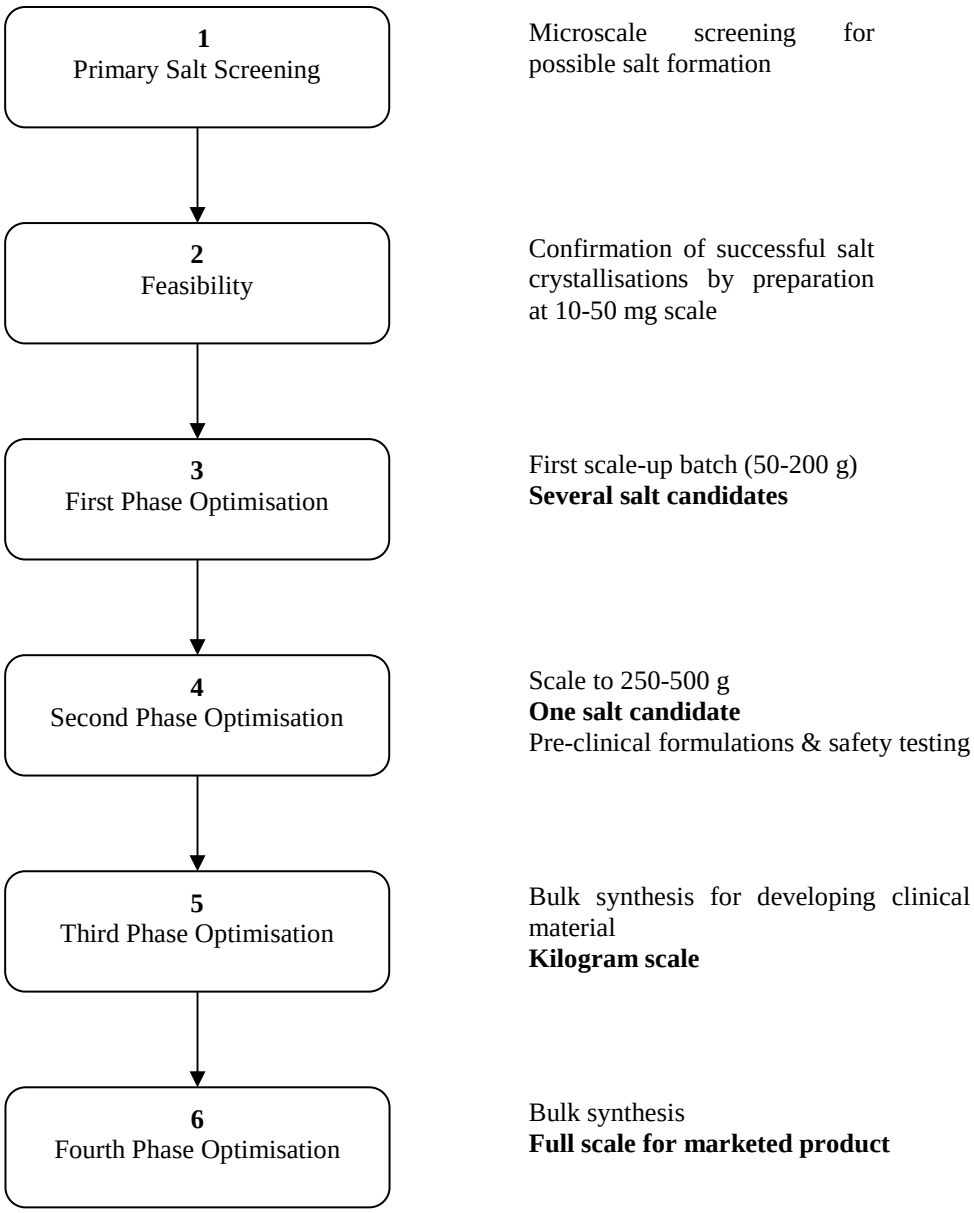
'Potentiometric cycling for polymorph creation' has also gained some interest within the crystal engineering community as a reliable method of screening for polymorphs of organic substances. The technique is an extension[60] of a potentiometric acid-base titration method described by Avdeef[61] and has led to the discovery of two polymorphic forms of sulindac[62], a non-steroidal anti-inflammatory drug, which also shows promise in anti-cancer treatment. When an ionisable compound such as sulindac dissolves in water, the pH of the solution will shift. Addition of small quantities of acid or base to this solution may change a supersaturated solution to a subsaturated one. Repeated cycling between these two states is possible by tiny adjustments to the pH of the solution. This is done by the addition of small quantities of acid or base. It is this 'chasing of equilibrium'[60] that allows the rapid measurement of the intrinsic thermodynamic solubility of the solute. The cycling is continued until there is substantial change in solubility of the solute. When such a change occurs, it is often an indication of a transition point and the appearance of a thermodynamically more stable polymorph in solution.

The milligram quantities of active pharmaceutical ingredients produced in industrial laboratories means that if solid state screening is to be effective, one has to carry out the crystallisation experiments on a much smaller scale than would otherwise be preferred. High throughput screening (HTS)[63] allows solid state chemists to conduct dozens of automated crystallisation experiments on a much smaller scale than is typically conducted in most crystallisation laboratories. Thus a large matrix of experimental variables can be investigated through HTS. Because of the limited material available, the so called 'gold standard' of solid state characterisation, X-ray powder diffraction, cannot be used to identify any new solid state forms that result from the screen, and instead Raman spectroscopy is commonly used.

1.4 Pharmaceutical salt selection

Pharmaceutical companies invest huge amounts of effort as well as capital on ensuring that the active pharmaceutical ingredients (API) of drugs maintain their desired physical properties upon scale up. A rigorous polymorph screen is now an integrated part of the drug development process. Provided the polymorph screen is performed at the early stages of scale up, the discovery of different polymorphic forms of the same substance provides an opportunity to market the drug in a form which best reproduces the desired physicochemical properties. If the API has undesired physicochemical properties and it is not possible to market it as an alternative polymorph with improved properties, other methods of modifying the physicochemical profile of the drug must be sought. In the pharmaceutical industry, salt formation is an established method for improving the solubility and hence bioavailability of drug candidates. In fact, an estimated half of all drugs used in medicinal therapy are administered as salts[64]. The term *salification* will be used to refer to the process of crystallising a salt from a given substance.

The *salt selection* process[65] refers to the discovery and development of the 'optimum salt form' (Scheme 1.8) of a drug substance. In this context, the optimal salt form refers to both chemical and physical properties. The choice of salt former (counter-ion) depends on various factors. From a preparative point of view, the primary factor is that it should yield a crystalline salt. This is because a highly crystalline salt allows purification and the removal of unwanted impurities. The absence of crystallinity is not a problem however if the intended dosage form is a liquid. The



Scheme 1.8: Illustrative flow diagram[66] for the industrial salt selection process of promising drug candidates.

choice of salt former is also limited by safety considerations, as dictated by the so called *GRAS*[67] (Generally Regarded As Safe) list of substances. For regulatory purposes (section 1.5), it is also important to know whether a drug is to be regarded as a *pharmaceutical equivalent* or a *pharmaceutical alternative*. Pharmaceutical equivalents contain the same amount of the active substance in the same dosage form, while pharmaceutical alternatives[68] are medicinal products that contain the same therapeutic

moiety but differ in the chemical form of that moiety or in the dosage form or strength. Salts are regarded as pharmaceutical alternatives.

For a given API, it is important to assess the extent to which the API is actually amenable to salt formation prior to initiating an expensive salt selection program (Scheme 1.8). Assuming the API contains at least one ionisable functional group, current industrial practice usually requires knowledge of two parameters prior to initiating a salt screen. These are the solution pK_a(s) and solubility of the API in a range of solvents. The pK_a of an ionisable group is the negative logarithm of the dissociation constant for that functional group. Provided a complementary set of pK_a(s) of the counterions to be screened are available, it is common to use a rule of thumb[69] which states that salt formation is likely when $\Delta pK_a \geq 3$. Although it is generally true that large values of ΔpK_a do increase the likelihood of salt formation, empirical evidence[70,71] suggests that there is no universal value of ΔpK_a at which salt formation is guaranteed for all acid-base pairs. This is a consequence of the fact that the solution pK_a and hence the extent of proton transfer is influenced by factors such as solvent, temperature and solution concentration, which are not considered in the current formulation of the ΔpK_a rule[69]. Thus trial and error still play an important role in the industrial salt selection process. Recent work by He *et al*[72] has shown that it is possible to predict multi-component crystal formation by using pulsed gradient spin echo ^{1}H NMR experiments to measure the solution intermolecular pair potentials of acid-base pairs. However, the method suffers from not being able to distinguish between a salt and cocrystal, without resorting to further ^{1}H NMR experiments to determine the likely atomic connectivity of the acidic proton. There are also question marks about the widescale applicability of this method given that it is a highly specialised method of experimentation and will require a fair amount of time to collect the data and subsequently analyse. The scarcity of a universally applicable predictive method for suggesting which acid-base pairs are likely to crystallise as salts, provide the motivation for validating computational modelling methods and this is covered in Chapters 4 and 5.

The solubility of the API in a given solvent is important for a couple of reasons. First of all, it allows us to assess the range of solvents that can be used to crystallise salts of the API. Secondly, knowledge of API solubility allows construction of a pH-solubility profile[66], which is used to quantify the pH range under which the salt is stable to dissociation. Of course from a practical point of view, the solubility of the API also

determines the method of drug delivery, including the size of the tablet if delivered in solid form.

Whilst for the most part, salification enhances the physicochemical properties of the parent API, there are a few negative aspects associated with the process that are worth mentioning here. The first of these relates to the percentage active content of the drug substance, which decreases markedly as counter-ions with higher molecular weights are used in the salt. The consequence of this is that a higher dose has to be administered to the patient if the free acid or base has moderate or low activity. Another potential problem that is exacerbated by salification is the increased tendency for the formation of solvates[73]. Section 3.3.3 of this thesis discusses in more detail the hydration behaviour of organic salts as a function of counterion.

1.5 Intellectual property issues surrounding salt selection

From the perspective of the pharmaceutical industry, performing an extensive salt selection program as early as possible is important for two reasons. First of all, it leads to the identification of the optimal solid form of the API before scale up and avoids the expensive and time consuming developmental studies inherent in late changes to the drug solid form. Equally as important, it allows the early patenting of the optimal salt form once it has been identified. By definition, patents give protection to inventions that solve a technical problem of some kind, provided the solution found is novel, and is susceptible to application by others in the field or industry. Once the patent has expired and assuming no extensions to the patent are granted, the invention may be used by others without fear of infringement. If many competitors are producing generic versions of previously patented drugs, this reduces the price and hence increases the number of people that can potentially benefit from the invention. Examples of common[74] types of patents are:

1. A *utility patent*. This has a life span of 20 years from the date of filing. It describes and claims any new and useful process, machine, article of manufacture, or composition of matter.
2. A *design patent*. This is valid for 14 years. It is for new, original and ornamental design for a manufactured article.

In the pharmaceutical industry, the patents of drugs fall under category 1. To be patentable, any invention must fulfil the requirements of *novelty, non-obviousness* and *usefulness*.

One of the tests for *novelty* is whether the proposed invention has been described in the "prior art". The prior art refers to the earlier work of others, or the inventor himself. Often controversies arise over this subject. The usual argument against the award of a patent is that the same chemical involved in the patent application is also found in the prior art.

An invention is *non-obvious*, if its "reduction to practice requires an inventive step so that it goes beyond the normal progress of technology and does not automatically derive from the state of the art"[64]. Often this requirement is met, if the invention could not have been made by a person of ordinary skill in the relevant field, without a reasonable expectation of success, using publicly available information and material.

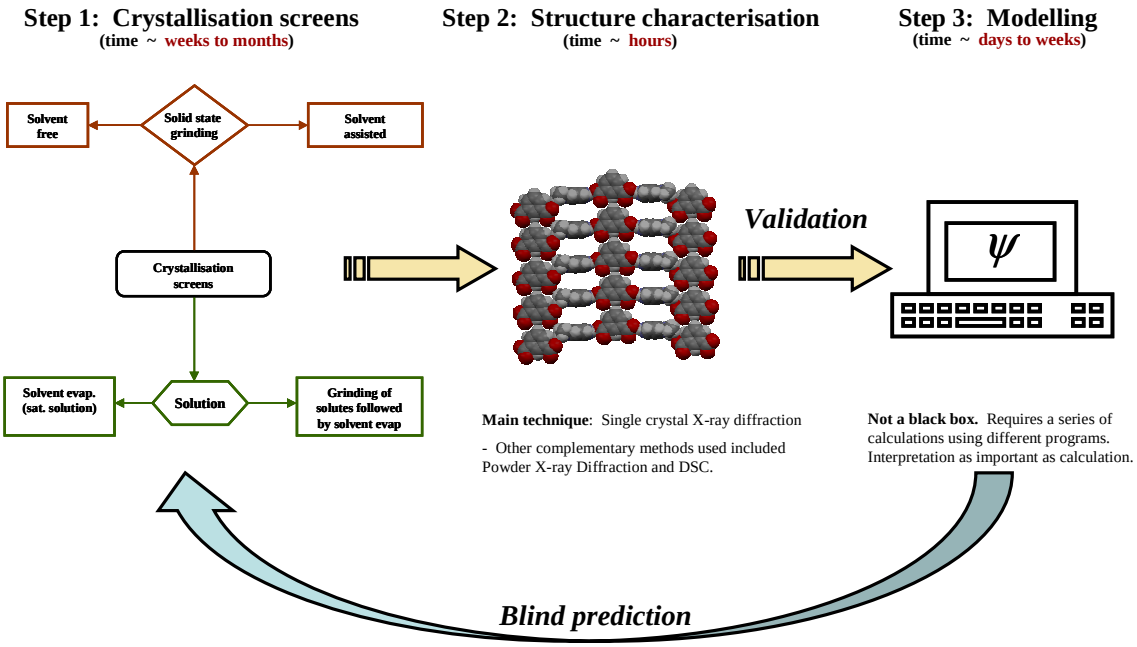
Finally for an invention to be deemed *useful*, it must have some practical utility other than just furthering research, at the time of filing the patent. In the context of pharmaceutical salt selection, it is not simply enough to discover a new salt form of a compound, especially if other salt forms are given in the prior art. Instead, the new salt must display some unexpected superior property when compared to the common salts that are already known. Obtaining patent protection for new salts often involves a so called *extension of term* to an additional patent that comprises the main compound in use. If an extension of term is sought, the inventor must prove the bioequivalence of the new salt and API. Once the patentability requirements outlined above have been met, patent law imposes requirements for invention disclosures. It is mandatory that in return for legal protection an inventor must provide an *enabling disclosure*. This is so as to allow others of ordinary skill in the art, to reproduce and use the product that is protected under the patent, without any difficult experimentation.

1.6 Crystal structure prediction as an aid to experimental salt screening

Computational crystal structure prediction refers to the *ab initio* prediction of the most energetically feasible crystal structures of a system using nothing but the chemical

diagram of the molecule(s) or ions as input. The practical motivation behind the development of a reliable method of crystal structure prediction (CSP) is 'to guide the design of new organic materials where the physical property of interest is very sensitive to the crystal structure'[75]. Small changes in the substituents of molecules, such as replacing a methyl group for an ethyl group, can lead to significant changes in the packing forces and hence the adopted crystal structure. Thus the ability to integrate a reliable method of CSP into the drug development program would mark a significant milestone in the targeted design of drugs and organic materials. Current methods of CSP have been extensively applied to neutral organic molecules where empirical[76] isotropic atom-atom potential parameters have been widely used to model the intermolecular forces between molecules. However, the majority of pharmaceutical drugs are marketed as multi-component solid forms[64], and salts in particular are widely used in industrial drug formulation. Over the past few years, there has been a gradual shift in the application of CSP from looking at the polymorphism of neutral organic molecules[77,78] to rationalising the crystallisation behaviour of multi-component solid forms[79,80]. This thesis is representative of this shift in emphasis and unlike the majority of the previous work on multi-component systems, will focus on the crystal structures of organic salts.

The majority of the work in this thesis will use CSP to rationalise the observed crystallisation behaviour of organic salts using the experimental structure and other available experimental information to validate the results of the computational model. Strictly speaking, the availability of experimental information makes such an endeavour an exercise in *post-diction* rather than *pre-diction.* This is highlighted in Scheme 1.9, which illustrates the typical approach adopted in this thesis. From left to right, the study begins (step 1) with a crystallisation screen for suitable crystal forms of the salt, using a range of experimental conditions. We are interested in good quality single crystals of the salt and a sufficiently extensive screen can take anywhere between weeks to months. If the experiments lead to good quality single crystals, structure characterisation is performed in step 2, using X-ray diffraction methods. It takes approximately 7 hours to collect a full sphere of reflection data for a good quality single crystal and whilst the actual diffraction experiment takes a short time, interpretation of the diffraction data can take longer for all but the simplest of structures. In step 3, the molecular structures of the ions are used as starting points for *ab initio* geometry optimisation, and the resulting



Scheme 1.9: From left to right, the orange arrows illustrate how CSP is applied in this thesis to rationalise the observed crystallisation behaviour of organic salts. The use of experimental information is important in validating the accuracy of the computational model. The blue arrow represents a truly blind prediction (Chapter 7), where the modelling is used to suggest the experimental crystal structure without any prior information.

gas phase optimised molecular conformations are used in a search for hypothetical crystal structures of the salt. When compared to the first two steps, the modelling generates a lot of data and it is not unusual to spend many weeks analysing the data and reaching relevant conclusions. Because the experimental structure is known, it can easily be located on the crystal energy landscape and often we would want to use the results of CSP to rationalise some observed experimental behaviour. Thus, the work presented in this thesis is important in validating whether existing methods of CSP can be used to complement experimental efforts in salt screening. Without the prior experimental information, this would not be possible as there would be no standard against which the theoretical results can be compared. It is hoped that future developments in the field will make it routine for this process to be done in reverse as experiments can be too costly and time consuming. This would use the results of CSP to postulate the likely experimental crystal structure(s). Chapter 7 of this thesis will assess the results of CSP when applied to the first blind prediction of an organic salt organised by the Cambridge Crystallographic Data Centre (CCDC).

Scheme 1.10: Illustration of two crystal energy landscapes, where each point represents a stable lattice energy minimum. Diagram (a) shows a crystal lattice energy landscape with one preferred structure (monomorphic) and diagram (b) shows a crystal energy landscape with many energetically competitive structures (polymorphic).

The output from a CSP study is a set of predicted lattice energy minima, depicted in a scatter plot of lattice energy *vs.* density (Scheme 1.10). This scatter plot is referred to as the crystal energy landscape and Scheme 1.10 shows two illustrative examples. One of the key results from previous work on the polymorphism of neutral organic molecules[50,77,81] has been the appearance of a crystal energy landscape with more energetically competitive crystal packings within a few kJ mol^{-1} than there are known polymorphs. This is illustrated in the so called[75] 'polymorphic' crystal energy landscape of Scheme 1.10. It is important to stress that the exact details of the crystal energy landscape are molecule dependent. Thus, not all organic molecules will have a 'polymorphic' crystal energy landscape as illustrated in Scheme 1.10. The alternative is that the system displays a crystal energy landscape with only one preferred structure in lattice energy (Scheme 1.10). Current evidence suggests that such 'monomorphic' crystal energy landscapes are much rarer for organic compounds but a notable example is provided by Pigment Yellow 74[82], whose experimental structure is 12 kJ mol^{-1} more stable than the nearest lattice energy minimum, in agreement with the results of experiments which have found only one crystal structure.

One explanation for the appearance of a 'polymorphic' crystal energy landscape for some organic molecules is that the 0 K lattice energy surface is producing too many minima as a result of the neglect of dynamic motions in the crystal lattice which could transform many of the observed lattice energy minima into the same free energy minimum. Indeed, a metadynamics[83] study on benzene[84] has shown only seven free energy minima corresponding to the known polymorphs, in contrast to the many

predicted lattice energy minima at 0 K[85]. The caveat that goes with these findings is the fact that benzene is atypical of many organic molecules as it lacks any hydrogen bond donor and acceptor groups. A similar metadynamics study on 5-fluorouracil[86], which has two polymorphic forms differing in the hydrogen bond motifs, has shown that 75 % of the predicted lattice energy minima are also free energy minima at ambient conditions of temperature and pressure. Thus, the free energy surface need not always be as simple as that observed for benzene, especially when there are one or more energetically competitive hydrogen bond motifs as a consequence of the hydrogen bond donor and acceptor groups on the organic molecule. It will be interesting to see the general complexity of salt crystal energy landscapes when compared to the picture that has emerged for neutral organic molecules. Certainly it is tempting to think that the dominant electrostatic contributions towards the lattice energies of salts will be more discriminating in the relative lattice energies of structures differing in the relative orientation of the hydrogen bonded ions, but this is a hypothesis that will be tested during the course of this thesis.

1.7 Outline and scope of this thesis

This thesis applies computational crystal structure prediction to organic salts with emphasis on chloride and carboxylate containing systems. Complementary experimental screening and characterisation is performed alongside the computational modelling for the majority of systems. The main method of structure characterisation used is single crystal X-ray diffraction.

Chapter 2 describes the theoretical basis for computational crystal structure prediction. It begins with a short summary of the various intermolecular forces at work in crystal structures and the assumptions made in modelling them. It then presents the numerical parameters of the atom-atom intermolecular potentials used to model repulsion-dispersion effects in the salts studied in this thesis and gives a critical assessment of their origin and transferability to charged systems. Finally, an existing method of crystal structure prediction based on the global minimisation of the lattice energy is presented. Chapter 3 presents the results of a literature review into the trends and properties of organic salt crystal structures as stored in the Cambridge Structural Database (CSD). The literature review showed that chlorides and carboxylates are the

two most frequently used counterions in salts derived from neutral organic bases and this was used to narrow the choice of salt systems studied in this thesis. Chapter 3 also presents the results of novel CSD surveys into the most prevalent proton acceptor in organic salts as well as the observed hydrogen bond motifs in carboxylate and chloride salts.

Chapter 4 presents the results of a combined experimental and computational study into the effects of acidic proton position in the crystal structures of organic salts. The chapter begins with the results of an extensive crystallisation screen for salts and cocrystals of pyridine and 4-dimethylaminopyridine using a range of dicarboxylic acids. The structures are characterised using a variety of experimental methods and these include single crystal and powder X-ray diffraction, differential scanning calorimetry and ^{1}H NMR spectroscopy. A selected set of solid forms from the experimental screens are used in computational modelling to test the sensitivity of salt crystal structures to the position of the acidic proton. Chapter 5 builds on this work and investigates the effects of proton position on the calculated crystal energy landscapes by using three model systems: a salt, a disordered salt-cocrystal system, and a cocrystal.

Chapter 6 extends the computational model of crystal structure prediction to chloride salts by looking at the pharmacologically active amantadine hydrochloride and memantine hydrochloride salts. All the experimental work presented in this chapter was performed by the group of Prof. Alistair Florence (Strathclyde University). The calculated crystal energy landscapes of the salts were used to rationalise their different polymorphic and hydration behaviours.

In Chapter 7, the results of a blind crystal structure prediction for 1,8-naphthyridinium fumarate are presented. The modelling was performed as part of the fifth international blind test of crystal structure prediction and provides a critical assessment of the success of the computational model used in this thesis. Finally Chapter 8 provides an overall conclusion to the work presented in this thesis and identifies areas that can benefit from further research.

2 Theoretical basis for crystal structure prediction of organic salts

2.1 Overview

This chapter describes the theoretical basis of computational crystal structure prediction as applied to organic salts. We begin with a short summary of the various intermolecular forces[87] at work in crystal structures and the assumptions made in modelling them. This will be followed by a section on model intermolecular potentials, where we discuss the origin of empirically derived dispersion-repulsion potentials and give the numerical values of the atom-atom potential parameters in the FIT(H_O,H_N) and W99 force fields used in this thesis. We will provide an outline of a method of computational crystal structure prediction based on the global minimisation of the lattice energy. This approach has been widely tested for rigid organic molecules and this thesis will extend the approach to the crystal structures of organic salts. Section 2.4 will summarise how CrystalPredictor[88] can be used to generate plausible packing arrangements of organic salts using a pre-defined dispersion-repulsion model and an electrostatic model comprised of potential derived atomic charges. Section 2.5 will present the crystal lattice simulation program, DMACRYS[89], and discuss its use in minimising the lattice energies of experimental or hypothetical structures using an accurate distributed multipole electrostatic model.

2.2 Physical origin of intermolecular forces

2.2.1 Introduction

Intermolecular forces are the forces that act between the molecules or ions of a given system. These forces are generally many orders of magnitude weaker than the forces that define the covalent bonds of the molecules or ions but are nevertheless important in simulations of the organic solid state because they determine the observed packing, density and morphologies of organic crystals. Often we talk about an intermolecular potential, *U(R)*, rather than an intermolecular force, *F(R)*, but the two are related according to the equation

$$F(R) = -\frac{dU(R)}{dR} \quad \textbf{(4)}$$

where R is the intermolecular distance between a pair of spherical molecules. For simulations of the organic solid state, the intermolecular force is also a function of the orientation, Ω, between the molecules or ions and so may act to reorient the molecules upon lattice energy minimisation. It is possible to classify the various intermolecular forces into one of two categories depending on the extent of overlap in the electron densities of a pair of molecules: *long range* effects where the interaction energy goes as the inverse power of the intermolecular distance, R, or *short range* effects where the interaction energy decreases exponentially with R. Long range effects consist of the *dispersion*, *electrostatic* and *induction* contributions while short range effects consist primarily of the *exchange-repulsion* contribution. The assumption of *pairwise additivity* is used to cut down on the computational cost associated with estimating the intermolecular potential. According to this, the intermolecular potential of a collection of molecules can be approximated as the sum of all two body interactions and we assume that the many body terms can be neglected because the series converges rapidly beyond the two body contributions. Not all intermolecular forces can be treated in this way however and the sections that follow will provide a brief outline of the physical origin of the dominant intermolecular forces as well as the assumptions made in modelling them.

2.2.2 Modelling electrostatic effects

The electrostatic energy plays a major role in determining the relative orientation and hence the observed hydrogen bond motifs between molecules/ions in a crystal. For two chemically distinct molecules, the electrostatic energy can be expressed as an integral over the charge densities of these molecules according to

$$U_{elec} = \int \frac{\rho^A(r_1)\rho^B(r_2)}{|r_1 - r_2|} d^3r_1 d^3r_2 \quad \textbf{(5)}$$

where $\rho^A(r_1)$ is the charge distribution of molecule A which is at a distance $|r_1 - r_2|$ away from molecule B. This definition of the electrostatic energy is implemented in Gavezzotti's SCDS-pixel method[34] which allows an estimate of the lattice energy for organic crystals by numerical integration over the *ab initio* charge density of the isolated molecule placed in the crystal structure. The mathematical formalism of Rayleigh-Schrödinger perturbation theory describes the electrostatic energy as the first order perturbation to the ground state energy of a pair of molecules/ions. According to this theory, the total interaction energy, U_{AB}, can be expanded as a series

$$U_{AB} = U_{AB}^0 + U_{AB}^1 + U_{AB}^2 + ... + U_{AB}^n \quad \textbf{(6)}$$

where U_{AB}^0 is the ground state energy derived by summing the monomer energies, E_A^0 and E_B^0. The first order correction to the total energy, U_{AB}^1, is then taken to be the electrostatic energy. Mathematically, we may define U_{AB}^0 and U_{AB}^1 as

$$U_{AB}^0 = \left\langle \psi_A^0 \mid H_A \mid \psi_A^0 \right\rangle + \left\langle \psi_B^0 \mid H_B \mid \psi_B^0 \right\rangle = E_A^0 + E_B^0 \quad \textbf{(7)}$$

and

$$U_{AB}^1 = \left\langle \psi_A^0 \psi_B^0 \mid V \mid \psi_A^0 \psi_B^0 \right\rangle = U_{elec} \quad \textbf{(8)}$$

where ψ_A^0 and ψ_B^0, are the groundstate wavefunctions of *A* and *B* respectively whilst H_A and H_B correspond to the Hamiltonian operators for the unperturbed *A* and *B* species. In equation 8, the perturbation operator, *V*, is defined as

$$V = \frac{1}{4\pi\varepsilon_0} \sum_{ij} \frac{e_i^A e_j^B}{r_{ij}} \quad \textbf{(9)}$$

where r_{ij} is the inter-particle distance between i and j whilst e_i^A and e_j^B represent the charges on the i^{th} and j^{th} particles of molecules A and B respectively. The term ε_0 represents the permittivity of vacuum.

One way of modelling the electrostatic contribution towards the intermolecular potential is to use a point charge model. This model assumes that the charge density of the isolated molecule can be described by the superposition of a set of spherical atomic electron densities. There are various methods for partitioning the total charge density into individual atomic contributions. One of the earliest efforts was made by Hirshfeld, who devised the so called "stockholder" partitioning method[90]. In the Hirshfeld stockholder partitioning scheme, the total electron density $\rho(r)$ is partitioned to individual atomic sites according to

$$\rho(r) = \sum_\alpha \rho_\alpha^H(r) \qquad \textbf{(10)}$$

where

$$\rho_\alpha^H(r) = \rho(r) d_\alpha^H(r) \qquad \textbf{(11)}$$

The terms $\rho(r)$ and $\rho_\alpha^H(r)$ are the total electron density and the fragment atomic contribution at a given site respectively, while the term $d_\alpha^H(r)$ represents the weight contribution of that atom to the total electron density. The method is so called because Hirshfeld used the stock market as an analogy and imagined that each atom participates locally in the molecular "profit" $\rho(r)$ in proportion to its share $d_\alpha^H(r)$ in the molecular "investment" $\rho^0(r)$. In this context the term $\rho^0(r) = \{\rho_\alpha^0(r)\}$ simply groups the densities of the free atoms giving rise to the reference electron density of the "promolecule"[91]. Another method is Bader's so called *Atoms in Molecules* approach[92], which uses a "zero flux" surface to divide the molecular charge density into a set of atomic charges. A zero flux surface is a two-dimensional surface on which the charge density is a minimum in the direction normal to the surface. For molecular systems, the charge density reaches a minimum between atoms. Thus the point at which the electron density is a minimum is taken as a natural place to partition the electron density into a

set of atomic volumes. Atomic properties such as the atomic charge and dipole moment can then be calculated by integrating the relevant operators over the atomic volume.

Most electronic structure modelling programs that calculate molecular charge densities, allow a method of fitting the numerical values of the atomic charges directly to the electrostatic potential around the molecule. These *potential derived charges*[93], which are used in the search algorithm, CrystalPredictor[88], have the advantage that they reproduce the electrostatic potential outside the van der Waals surface of the molecule.

An alternative to the simple point charge model described above is the distributed multipole model, where the electron density is described by a series of individual multipole moments that are centred at specified sites. These sites are often taken to be the atomic nuclei. Distributed multipoles can be derived from Bader partitioning but the method leads to multipole expansions that are only slowly convergent. The method used in this thesis is Stone's Distributed Multipole Analysis (DMA)[94,95], which starts from the one electron density matrix of the *ab initio* wavefunction. The wavefunction is represented in terms of a set of Gaussian orbitals. The DMA then uses the properties of Gaussian orbitals to calculate the multipole moments at a number of specified sites. The main strength of DMA when compared to other methods for partitioning the total electron density is the relatively fast rate of convergence for the multipole moments. Because distributed multipoles can represent the anisotropy of the atomic charge distribution within a molecule, the electrostatic potential arising from key bonding features such as lone pairs and π-electron clouds are automatically reproduced. For a set of 50 rigid molecules, Day[96] has contrasted the results of crystal structure prediction using an atom centred multipole model versus a point charge model for the electrostatic contribution to the lattice energy. He found that on the whole, an electrostatic model consisting of atom centred multipoles is more successful in reproducing the experimental structure at or close to the predicted global minimum in lattice energy than is the case with a point charge model. For the work presented in this thesis, atomic multipoles up to rank 4 (hexadecapole) will be used.

2.2.3 Induction energy

Induction effects are not explicitly modelled in this thesis but the empirically derived atom-atom intermolecular potentials used will have absorbed the effects of induction during the fitting to organic crystal structures. According to equation 6, which defines the total intermolecular energy between a pair of molecules, the second order correction to the total energy describes the induction and dispersion contributions. Rayleigh-Schrödinger perturbation theory defines the second order correction to the energy as

$$U_{AB}^{2} = - \sum_{\substack{\psi_A^m \neq \psi_A^0 \\ \psi_B^n \neq \psi_B^0}} \frac{\left\langle \psi_A^0 \psi_B^0 \left| V \right| \psi_A^m \psi_B^n \right\rangle \left\langle \psi_A^m \psi_B^n \left| V \right| \psi_A^0 \psi_B^0 \right\rangle}{U_{AB}^{mn} - U_{AB}^{0}} = U_{induction}^{A} + U_{induction}^{B} + U_{dispersion}^{AB} \quad \textbf{(12)}$$

and the summation in the second order correction to the energy is explicitly over excited states. The explicit summation over excited states means that the contribution where both molecules are in the ground state (m=n=0) is neglected. Thus we consider terms in the summation that arise due to molecule *A* being in an excited state while molecule *B* is in the ground state ($U_{induction}^{A}$), or the configuration where molecule *B* is excited while molecule *A* is in the ground state ($U_{induction}^{B}$) or finally the configuration where both molecules *A* and *B* are in an excited state ($U_{dispersion}^{AB}$). Thus the first two configurations model the induction contribution towards the second order correction whilst the final configuration models the dispersion contribution towards the second order energy. Both the induction and dispersion contributions are attractive but only the dispersion contribution is approximately pairwise additive.

Reliable methods of estimating induction effects are computationally demanding because they require the use of large basis sets and high quality wavefunctions in order for the polarisability tensors to converge. For a given direction of the electric field, the polarisability tensor defines the extent to which the molecule is polarised in the *x*, *y*, and *z* directions. The modelling of the induction contribution towards the lattice energy is important for simple ionic salts because of the presence of charged species and the influence such species have in polarising the charged clouds of neighbouring molecules or ions. Moreover, there is growing experimental evidence which shows that induction effects play a significant role in the crystal structures of even non-polar organic

molecules such as naphthalene[97]. Welch[98] *et al* contrasted two methods of estimating the induction energy in organic crystal structures. The first was a distributed atomic polarisability model combined with distributed multipole moments. The second involved the *ab initio* calculation of the molecular charge density in a point charge field. In both methods, the magnitude of the induction contribution towards the lattice energy was estimated by calculating the induced moments in a molecule that is located at the centre of a large cluster that extends to at least 15 Å. This cluster was assumed to represent the extended crystal structure. In both cases, the induced moments were iterated to self-consistency in the cluster before using them to evaluate the induction contribution towards the lattice energy of the crystal. Welch found that the two models are in agreement with one another, despite the differing assumptions made in each. Where the hydrogen bonding motifs of two polymorphic structures are very different, he also found that the induction energy contribution can be important in re-ordering the relative lattice energies of the structures. This thesis will not be modelling the induction contribution towards the lattice energies of salts explicitly as the majority of the research was performed prior to the recent implementation of induction into the crystal lattice simulation program, DMACRYS[89].

2.2.4 Dispersion energy

Dispersion forces have a purely quantum mechanical origin and arise from the instantaneous correlation of fluctuations in the charge densities of molecules or ions. In other words, the zero point motion of the electrons on one molecule creates a temporary dipole which induces a correlated dipole in a neighbouring molecule. These forces are universally attractive and they are the reason why non-polar molecules condense to form liquids. In the context of perturbation theory, the dispersion energy can be expressed mathematically as follows

$$U_{dispersion}^{AB} = -\sum_{\substack{\psi_A^m \neq \psi_A^0 \\ \psi_B^n \neq \psi_B^0}} \frac{\left\langle \psi_A^0 \psi_B^0 \,|\, V \,|\, \psi_A^m \psi_B^n \right\rangle \left\langle \psi_A^m \psi_B^n \,|\, V \,|\, \psi_A^0 \psi_B^0 \right\rangle}{U_B^n - U_B^0 + (U_A^m - U_A^0)} \qquad \textbf{(13)}$$

where the various terms are the same as in equation 12. The most widely understood treatment of dispersion forces is that provided by London[99] where atoms can be visualised as a set of harmonic oscillators. In this treatment, it is the coupling of the oscillators which leads to the lowering in energy. The energy depends on the square of the coupling constant. For a pair of interacting dipoles, this coupling is proportional to R^{-3}, so the leading term in the dispersion energy is proportional to R^{-6}.

2.2.5 Exchange-repulsion energy

The exchange-repulsion energy is the net repulsive contribution towards the total intermolecular energy. Both the exchange and repulsion components act at short range but they have opposite signs. When the wavefunctions of two molecules overlap, their charge densities repel one another. This repulsion contribution towards the intermolecular energy arises from the Pauli exclusion principle, which forbids electrons of the same spin occupying the same space. In addition to this dominant repulsion contribution, the indistinguishable nature of electrons means that they are able to exchange between the orbitals of different molecules. This leads to a weak attractive force. In this thesis, the exchange-repulsion energy will be modelled using a single exponential function and further details can be found in the next section.

2.3 Model intermolecular pair potentials

This thesis will model the crystal lattice energies of organic salts using an empirically derived repulsion-dispersion model coupled with an electrostatic model derived from a distributed multipole analysis of the *ab initio* wavefunctions of the isolated ions. The repulsion-dispersion model takes the form of a Buckingham *exp*-6 function which may be written as

$$U_{rep-disp}^{MN} = \sum_{i\in M, k\in N} A_{\iota\kappa} \exp(-B_{\iota\kappa} R_{ik}) - \frac{C_{\iota\kappa}}{R_{ik}^6} \qquad \textbf{(14)}$$

where the interactions are between atom *i* of type ι in molecule *M* and atom *k* of type κ in molecule *N*. The exponential term represents the repulsive contributions that set in at short distances while the negative R^{-6} term models the long range dispersion

contributions. In using equation 14, we assume that the interactions between the molecules or ions can be approximated by the interactions between their constituent atoms. We also assume that the empirical atom-atom parameters (A_{ik}, B_{ik} and C_{ik}) are transferable between different molecules. The parameters A_{ik}, B_{ik} and C_{ik} are specific to a pair of interacting non-bonded atoms and determine the strength of the repulsion-dispersion forces as a function of the intermolecular separation, R_{ik}. Provided the functional groups on the system to be modelled are the same as those found in the molecules used to parameterise the intermolecular potentials, transferability is a reasonable assumption to make. However, the assumption can lead to erroneous results in certain cases where the crystal structure displays an unusual intermolecular interaction or van der Waals contact distance that is not sampled in the fitting of the intermolecular potential parameters. The parameters A_{ik}, B_{ik} and C_{ik} are obtained by fitting to the crystal structures and heats of sublimations of organic materials. The numerical values of these heteroatomic parameters can be calculated from the homoatomic parameters using the following combining rules.

$$\begin{aligned} A_{ik} &= \sqrt{A_{ii} A_{kk}} \\ B_{ik} &= \frac{1}{2}\left[B_{ii} + B_{kk}\right] \\ C_{ik} &= \sqrt{C_{ii} C_{kk}} \end{aligned} \qquad \textbf{(15)}$$

In the work presented in this thesis, the atom-atom potential parameters used will be taken from either the FIT or W99 force fields (Table 2.1). The FIT potential parameters were derived by Williams and co-workers who initially fitted the C and H (bonded to C or H_C) parameters to a set of hydrocarbon[100] crystal structures. Subsequent work by Williams led to the derivation of chlorine[101], oxygen[102], nitrogen[103] and fluorine[104] potentials. The potential parameters for C and H_C quoted in Table 2.1 are those derived from Williams work on azahydrocarbons[103]. Given that many organic crystal structures display hydrogen bonds with one or more polar X-H (X=N or O) bonds, it was necessary to distinguish between hydrogens attached to C and polar hydrogens attached to N or O atoms. This led Coombes[105] to derive a potential for protons attached to nitrogen (referred to as H_N in Table 2.1) by fitting to the crystal structures of 40 rigid

Potential	Atom pair	Description	A / kJ mol^{-1}	B / Å^{-1}	C / kJ mol^{-1} Å^{6}
FIT	C···C	Any C atom	3697463	3.60	2439.82
W99	C(2)···C(2)	C bonded to 2 atoms	103235	3.60	1435.09
W99	C(3)···C(3)	C bonded to 3 atoms	270363	3.60	1701.73
W99	C(4)···C(4)	C bonded to 4 atoms	131571	3.60	97836
FIT	H···H	H bonded to C	11971.1	3.74	136.4
W99	H(1)···H(1)	H* bonded to C	12680	3.56	278.37
FIT	H_O···H_O	H bonded to O	2263.3	4.66	21.5
W99	H(2)···H(2)	H* in alcoholic group	361.3	3.56	0
W99	H(3)···H(3)	H* in carboxyl group	115.7	3.56	0
FIT	H_N···H_N	H bonded to N	5029.68	4.66	21.5
W99	H(4)···H(4)	H* bonded to N	764.9	3.56	0
FIT	N···N	Any N atom	254531.05	3.78	1378.41
W99	N(1)···N(1)	N in triple bond	96349	3.48	1407.57
W99	N(2)···N(2)	N (not triple bonded) and no H	102369	3.48	1398.15
W99	N(3)···N(3)	N bonded to 1 H atom	191935	3.48	2376.55
W99	N(4)···N(4)	N with 2 or more bonded Hs	405341	348	5629.82
FIT	O···O	Any O atom	230064	3.96	1123.6
W99	O(1)···O(1)	O bonded to 1 other atom	241042	3.96	1260.73
W99	O(2)···O(2)	O bonded to 2 other atoms	284623	396	1285.87
FIT	F···F	Any F atom	363725	4.16	844
FIT	Cl···Cl	Any Cl atom	924675	3.51	7740.48

Table 2.1: FIT and W99 *exp*-6 dispersion-repulsion parameters used in this thesis. * A key feature of the W99 force field is that hydrogen interaction sites are moved by 0.1 Å into the H-X bond from their neutron or *ab initio* optimised positions.

organic molecules, the majority of which contained the polar N-H bond. Beyer[106] later found that scaling the pre-exponential repulsion parameter of H_N by 0.45 leads to a potential which is suitable for modelling polar protons attached to oxygen atoms (referred to as H_O in Table 2.1) in carboxylic acids. In this thesis, when the H_O and H_N potentials are used in conjunction with Williams' original FIT potential parameters, the resulting force field will be referred to as FIT(H_O,H_N). Similarly if the crystal structure contains only one polar proton of the kind H_O or H_N, the resulting force fields will be referred to as FIT(H_O) or FIT(H_N) respectively.

The W99[107] force field contains parameters for C, H, N and O, which depend on the hybridisation state and atomic connectivity. It was validated against a range of organic crystal structures including peptides and nucleosides. In deriving the intermolecular atom-atom potentials of the W99 force field, the hydrogen interaction sites were foreshortened by 0.1 Å into the H-X bond from the neutron or *ab initio* optimised positions. This foreshortening reflects the significant shift[100] in the apparent centre of electron density towards the non-hydrogen atom in H-X bonds.

Because the potential parameters in the FIT and W99 force fields were fitted to many organic crystal structures, the effects of induction were absorbed into the potential parameters in an average way. Furthermore, the potential parameters in the FIT and W99 force fields were fitted to the crystal structures of neutral organic molecules and not the crystal structures of organic salts. Given these limitations in the available force fields, this thesis will test the sensitivity of the experimental structures to the FIT and W99 force fields prior to any crystal structure prediction studies. This will be done by computing the lattice energy minimum of the experimental structure using the chosen force field in conjunction with a distributed multipole electrostatic model. In this thesis, two types of lattice energy minima will be calculated when validating the intermolecular potential against the experimentally determined structure. The first represents the experimental structure minimised with respect to the experimental conformations and this is abbreviated as the *ExpMinExp*. Prior to calculating the *ExpMinExp*, the bondlengths to hydrogen atoms are elongated to standard neutron values and this counteracts the foreshortening of H-X bonds introduced by the X-ray diffraction experiment. The second type of lattice energy minimum is calculated with *ab initio* optimised molecular conformations for the ions in the salt and is abbreviated as the *ExpMinOpt*. The structural deviations between the experimental and *ExpMinExp* structures provide information on the sensitivity of the structure to the force field used. A suitable force field is expected to reproduce the intermolecular distances to within the likely[108] errors of neglecting the thermal effects. Significant differences in the *ExpMinExp* and *ExpMinOpt* structures usually indicate differences in the molecular conformations of the ions in the two structures. In this thesis, the crystal structure prediction of salts will be performed using the same *ab initio* optimised molecular conformations of the ions used to calculate the *ExpMinOpt*. This means that the *ExpMinOpt* is the closest match to the experimental structure one could hope to find in the lattice energy landscape and this allows for easy identification of the experimental structure.

2.4 Using CrystalPredictor to generate hypothetical crystal structures

2.4.1 Introduction

A crystal structure is defined by six lattice parameters (*a*, *b*, *c*, α, β, γ), a space group and the positions of all the atoms in the crystallographic asymmetric unit. The asymmetric unit defines the smallest possible repeat unit from which the whole crystal structure can be built by applying space group symmetry operations. The cell volume and cell density are derived from knowledge of these six lattice parameters. For a given system, the aim of any crystal structure prediction algorithm is to calculate the range of energetically feasible crystal packings in the absence of any prior knowledge of the unit cell. This problem of generating plausible lattice energy minima on the potential energy surface, has been termed the *packing problem*[109]. The search space considered must be large enough so as to encompass possible polymorphic forms of the global minimum structure. There are two contrasting computational modelling approaches to solving the packing problem. The first has a statistical basis, and relies on analysing the crystal structures held in the Cambridge Structural Database (CSD[29]) for common packing types. These packing types are used to define suitable crystallographic descriptors (*i.e.* co-ordination types and common space groups), which are either built-in to the software used to generate the hypothetical crystal structures (MOLPAK[110]) or are user defined during the crystal structure generation step. The purpose of using such statistical data is to narrow the search space considered in generating the hypothetical crystal structures. The second approach, and the one used in this thesis, is a purely mathematical one that uses no crystallographic insights from previously published structures stored in the CSD. The program CrystalPredictor[88] serves as a good example of a search algorithm that falls in this latter category and the following section provides more details about how it works. In this thesis, CrystalPredictor was used to calculate the crystal energy landscapes of salts because unlike MOLPAK, the CrystalPredictor search algorithm is capable of searching phase space with two or more chemically independent molecules/ions in the asymmetric unit.

2.4.2 Rigid body searches for hypothetical crystal structures

CrystalPredictor implements a global optimisation search algorithm. Three steps can be identified in this search algorithm:

1. Global search for plausible crystal packing arrangements
2. Local lattice energy minimisation of candidate structures under space group symmetry constraints
3. Post-processing of predicted structures to identify unique lattice energy minima

In the global search step, CrystalPredictor begins by specifying the space group of the candidate structure. By default, the program generates hypothetical crystal structures in one of the 59 space groups with a frequency of more than 0.05 % for organic molecules in the Cambridge Structural Database[111] (CSD). In principle, the global optimisation search algorithm used by CrystalPredictor is capable of predicting crystal structures in each of the 230 possible space groups. In practice however, organic molecules rarely crystallise in space groups other than the most common triclinic, monoclinic and orthorhombic crystal systems. Hence, the default 59 space groups considered by CrystalPredictor. The probability of generating a candidate structure in a given space group *s* is proportional to $N_s(F_s)^{0.75}$ where N_s is the number of symmetry relations in *s* and F_s is the number of organic molecules known to crystallise in that space group at the time the program was written[88]. Once the space group is specified, CrystalPredictor determines the numerical values for a set of *decision variables*, which are used to determine the geometry of the unit cell and positions of molecules within it. The decision variables optimised are

1. The lattice lengths $l_a \in \left(l^L, l^U\right), a \in \{1, 2, 3\}$
2. The lattice angles $\omega_a \in \left(\omega^L, \omega^U\right), a \in \{1, 2, 3\}$
3. Normalised positions of the molecular centres of mass $\hat{\boldsymbol{r}}_j$, $j = 2, ..., \mathrm{Z}$ of all species in the unit cell
4. The orientations of the Z molecules in the unit cell

The superscripts L and U define the lower and upper bounds for the lattice parameters, which are a function of the space group. It is possible to define a cell matrix according to $\boldsymbol{R} = \boldsymbol{R}\left(l_a, \omega_a, a = 1, 2, 3\right)$. In the crystal structure generation step, a large number of random starting values (within the given bounds) are considered for the numerical

contents of the cell matrix, $\boldsymbol{R}$. There are a number of ways of generating such starting points. One approach is to carry out a Monte Carlo sampling, which relies on uniformly distributed pseudo-random numbers[112]. The disadvantage with the Monte Carlo sampling is that it offers no guarantee of uniform coverage for the sampling limits considered. One can overcome this difficulty by using low discrepancy sequences, such as the one proposed by Sobol[113]. Sobol sequences, as implemented in CrystalPredictor, have been shown to give a more uniform coverage when compared to that obtained using a random sequence.

Once all decision variables have been assigned values that are subject to space group symmetry constraints, the resulting cell is checked to see whether it is sensible or should be excluded from further consideration. Structures are eliminated at this stage if the crystal density is below a given threshold or the lattice energy exceeds a specified threshold. Also, structures with intermolecular atom-atom distances of less than 0.5 Å are eliminated at this stage. Structures that pass these tests are used as starting points for local minimisation of the lattice energy. The crystal lattice energies of hypothetical structures are estimated using a Buckingham *exp*-6 potential and an electrostatic model that consists of atomic charges derived by fitting to the electrostatic potential of the *ab initio* single molecule wavefunction(s).

The final step is to identify which of the predicted lattice energy minima are unique. CrystalPredictor compares the lattice energies (using an atomic point charge model), cell densities and intermolecular distances between atoms in any pair of structures to be compared. Because the lattice energies estimated by CrystalPredictor use an atomic point charge model for the electrostatic interactions and not an accurate distributed multipole model, it is necessary to re-minimise the lattice energies of the unique structures generated by CrystalPredictor. Section 2.5 describes how this can be done using DMACRYS[89].

2.4.3 Comparing the experimental and predicted lattice energy minima

One of the strengths of computational crystal structure prediction is that it generates thousands of plausible crystal structures which can be used to gain a deeper understanding of the observed crystallisation behaviour. Not all of the predicted low energy structures will be observed experimentally and this reflects the inaccuracies in

the model for the energies as well as the neglect of kinetic effects in the predictions. Methods of comparing the experimental and calculated lattice energy minima are important because such methods provide confidence that the intermolecular potential used reproduces the experimental structure within the accuracy of neglecting the thermal effects. In this thesis, the success of reproducing the experimental structure in the *ExpMinExp* or *ExpMinOpt* will be quantified by the *F*-value[76]

$$F = \left(\frac{100\Delta a}{a}\right)^2 + \left(\frac{100\Delta b}{b}\right)^2 + \left(\frac{100\Delta c}{c}\right)^2 + \Delta\alpha^2 + \Delta\beta^2 + \Delta\gamma^2 + (10\Delta x)^2 + \left(\frac{\Delta\theta}{2}\right)^2 \quad \textbf{(16)}$$

where Δx is the total rigid-body translational displacement, $\Delta\theta$ is the total rigid-body rotational displacement and the other terms describe changes in the cell lengths and angles. The quantity *F* is sometimes referred to as the 'figure-of-shame' and describes the accuracy in reproducing the experimental structure given the assumed molecular model and intermolecular potential used in the minimisation. The *F* value is typically an order of magnitude larger in calculations leading to the *ExpMinOpt* when compared to lattice energy minimisations leading to the *ExpMinExp*. This is a consequence of the differences in the molecular conformation(s) between the experimental and *ExpMinOpt* structures. Such differences in conformation introduce errors in the optimal hydrogen bond geometries and hence lead to molecular re-arrangements that contribute to a larger *F*-value after lattice energy minimisation.

The program Mercury[114] is extensively used in this thesis to visualise the packings of the experimental or predicted crystal structures. The program will also be used to quantify the similarities in crystal packing for pairs of structures. Such crystal packing similarity calculations are useful for comparing the experimental structure with the hundreds of predicted structures that are usually within 20 kJ mol^{-1} of the global minimum or within the natural cut-off of the energy landscape. The calculations are based on the program COMPACK[115], which defines a cluster of molecules in the reference and comparison structures and then quantifies the geometrical difference in the clusters based on a root mean square deviation. As part of these calculations, Mercury overlays the two structures by superimposing regions of common packing. The default size for the clusters is 15 molecules and the calculated root mean square deviation is referred to as $RMSD_{15}$. The size of the cluster is an adjustable parameter as

are the default tolerances of 20 % deviation in the intermolecular distances and 20 ° deviation in the angles between molecular fragments. The similarity index of the powder patterns for any pair of structures is calculated and this gives further confidence in quantifying the similarities of structures. In addition to the above applications of Mercury, the program will be used to calculate the graph sets of hydrogen bond motifs.

2.5 Using DMACRYS to estimate crystal lattice energies

2.5.1 Introduction

DMACRYS[89] estimates the lattice energies (defined in section 1.2.3) of organic crystals using an accurate distributed multipole electrostatic model coupled with an empirically derived atom-atom dispersion-repulsion model. DMACRYS has been developed with the modelling of multi-component solid forms in mind and as such can model structures with large organic molecules or multiple molecules/ions in the crystallographic asymmetric unit. In the following sections, we will discuss how DMACRYS can be used to estimate the lattice energies of organic salts. First we will present the mathematical basis for minimising the crystal lattice energy assuming no change in the molecular conformations during structure relaxation. This so called *rigid body* approximation applies in simulations of crystals with little or no torsional flexibility for the molecules/ions. Then we discuss how DMACRYS can be coupled with CrystalOptimizer[116] to model the distortion of the molecular/ionic conformations induced by the crystal packing forces.

2.5.2 Estimating the lattice energies of salts composed of rigid ions

The static lattice energy minimisation code, DMACRYS, implements a multipolar description of the electrostatic potential surrounding the reference ions. It takes as input a SHELX or FDAT file that contains the lattice parameters, atomic positions expressed as fractional co-ordinates and the symmetry cards unique to the crystal space group. The input files for DMACRYS are generated by the utility program, NEIGHCRYS. Before NEIGHCRYS can generate the necessary input files to DMACRYS, the user must provide two pieces of information. The first is a list of maximum atom-atom

covalent bond distances expected in the chemically distinct molecules/ions of the asymmetric unit. This allows NEIGHCRYS to calculate the bond connectivity and hence define the rigid ions. The second piece of information is an axis system for each chemically distinct ion in the asymmetric unit. Only three atoms are needed to define an axis system for each molecular ion, two defining an axis direction and the third defining a plane. Usually the atoms defining the axis system are chosen to reflect any internal molecular symmetry. If these two pieces of information are available, NEIGHCRYS will read the input SHELX file for the cell parameters, symmetry cards and fractional atomic co-ordinates of the crystal and convert the cell into an orthonormal Cartesian axis system with the *c*-axis usually oriented along Z. Unlike the molecule-fixed axis system described earlier, this global axis system for the unit cell is internal to NEIGHCRYS.

DMACRYS estimates the second derivatives, torques and non-central forces that arise from the anisotropy of the intermolecular potential due to the multipolar interactions. These forces and torques are transferred to the centre of mass of the ions and used to determine the strains on the rigid ions within the crystal lattice. The strains on the rigid ions determine the translation and rotation of the rigid ions during lattice energy minimisation. It is possible to define the change in the crystal structure as a vector, $\boldsymbol{\delta}$, whose components are the three translation and rotation vectors per molecule as well as six strain matrix elements. Thus, the intermolecular lattice energy as a function of small changes in the lattice parameters can be expressed as

$$U_{inter}(\mathbf{r}') = U_{inter}(\mathbf{r}) + \boldsymbol{\delta}^T \bullet \mathbf{g} + \frac{1}{2}\boldsymbol{\delta}^T \bullet \mathbf{W} \bullet \boldsymbol{\delta} \qquad \textbf{(17)}$$

where **g** is a vector of first derivatives and $\boldsymbol{W}$ is a matrix of second derivatives (or Hessian)[117]. From the calculated $\boldsymbol{g}$ and $\boldsymbol{W}$, DMACRYS defines the displacement from equilibrium as

$$\boldsymbol{\delta} = -\frac{\boldsymbol{g}}{\boldsymbol{W}} \qquad \textbf{(18)}$$

and uses a modified Newton-Raphson search method to converge on the lattice energy minimum. The Hessian, $\boldsymbol{W}$, determines whether the calculated lattice energy minimum

is a saddle point or at a true minimum on the potential energy surface. A saddle point is characterised by a Hessian with at least one negative eigenvalue. If the minimisation reaches a saddle point, DMACRYS can remove the negative eigenvalue that corresponds to this saddle point and relax the crystal structure to a symmetry sub-group of the original crystal symmetry. The sub-group will be lower in energy when compared to the saddle point. Each symmetry representation that is removed by DMACRYS will lead to a doubling of the contents of the asymmetric unit.

When estimating the crystal lattice energy, DMACRYS computes the dispersion-repulsion contributions towards the lattice energy by summing over all atom-atom intermolecular interactions up to a pre-defined cut-off. The default value is 15 Å but this may be modified depending on the size of the rigid ions. For the electrostatic energy, the Ewald summation technique[118] is used to sum the charge-charge (R_{ab}^{-1}), charge-dipole (R_{ab}^{-2}) and dipole-dipole (R_{ab}^{-3}) terms with all other multipole terms up to R_{ab}^{-5} summed to a 15 Å direct space cut-off between the centres of mass of the molecules/ions.

2.5.3 Estimating the lattice energies of salts composed of flexible ions

When an organic salt consists of ions with one or more flexible degrees of freedom, it is no longer possible to treat the ions as rigid bodies during lattice energy minimisation. Instead, the relevant flexible degrees of freedom must be optimised as part of the lattice energy minimisation. In this thesis, CrystalOptimizer[116] will be used to refine the lattice energies of salts containing flexible ions according to the equation $E^{latt} = E^{inter} + \Delta E^{intra}$. Changes in the intramolecular energies (ΔE^{intra}) are estimated *ab initio* using GAUSSIAN03[119] and for a particular conformation, the intermolecular energy (E^{inter}) is estimated with DMACRYS. Empirical force fields[120] will not be used to estimate the ΔE^{intra} associated with molecular deformation, because such force fields are often derived by treating the inter- and intramolecular components separately[121] and this can lead to non-physical distortions in the molecular conformations. In principle, it is possible to refine any number of degrees of freedom (bondlengths, bond angles and torsions) with CrystalOptimizer. In practice however, it is sufficient to refine only a small number of torsion angles as this captures most of the flexibility exhibited by the

molecule/ion. Thus the majority of the intramolecular degrees of freedom ($\boldsymbol{\theta}$) are not expected to change from their gas phase optimised values. We define these intramolecular degrees of freedom as *rigid* ($\boldsymbol{\theta}^r$) and they are distinguished from the remaining *flexible* ($\boldsymbol{\theta}^f$) intramolecular degrees of freedom that are expected to change under the effects of the crystal packing forces. For a particular input conformation, DMACRYS computes analytically, the gradients of the lattice energy with respect to the lattice variables ($\boldsymbol{X}$). This allows CrystalOptimizer to re-formulate[116] the equation for the lattice energy

$$\min_{\boldsymbol{X},\boldsymbol{\theta}^f} E^{latt} = \min_{\boldsymbol{X},\boldsymbol{\theta}^f}\left\{\Delta E^{intra}\left(\boldsymbol{\theta}^f\right) + E^{inter}\left(\boldsymbol{X},\boldsymbol{\theta}^f;\boldsymbol{\theta}^r\right)\right\} \quad \textbf{(19)}$$

as a two-level optimisation problem

$$\min_{\boldsymbol{\theta}^f} E^{latt} = \min_{\boldsymbol{\theta}^f}\left\{\Delta E^{intra}\left(\boldsymbol{\theta}^f\right) + \min_{X} E^{inter}\left(\boldsymbol{X},\boldsymbol{\theta}\right)\right\} \quad \textbf{(20)}$$

where

$$\Delta E^{intra}\left(\boldsymbol{\theta}^f\right) = \min_{\boldsymbol{\theta}^r}\left\{E^{intra}\left(\boldsymbol{\theta}^r;\boldsymbol{\theta}^f\right)\right\} - E^{vac} \quad \textbf{(21)}$$

and

$$\boldsymbol{\theta}^r\left(\boldsymbol{\theta}^f\right) = \arg\min_{\boldsymbol{\theta}^r}\left\{E^{intra}\left(\boldsymbol{\theta}^r;\boldsymbol{\theta}^f\right)\right\} \quad \textbf{(22)}$$

In the outer optimisation, the values of the flexible degrees of freedom, $\boldsymbol{\theta}^f$, are varied. For a particular set of $\boldsymbol{\theta}^f$, CrystalOptimizer calls GAUSSIAN to perform a quantum mechanical minimisation of the intramolecular energy. This leads to values for the rigid degrees of freedom, $\boldsymbol{\theta}^r$, and the corresponding intramolecular energy penalty, ΔE^{intra}, for changes in $\boldsymbol{\theta}^f$. In the inner optimisation, DMACRYS is used to compute the minimum intermolecular energy and the values of the lattice variables, $\boldsymbol{X}$, for a given set of input molecular conformations ($\boldsymbol{\theta}^f$, $\boldsymbol{\theta}^r$).

If each outer optimisation step is performed *ab initio*, accurate estimates of the deformation energies, ΔE^{intra}, would be obtained but this would be at the expense of the

high computational cost associated with these quantum mechanical calculations. CrystalOptimizer reduces the computational cost of these outer optimisations without compromising on the accuracy of the calculations. It does this by estimating ΔE^{intra} using a Local Approximate Model (LAM) based on a quadratic Taylor expansion constructed about a reference point, $\boldsymbol{\theta_{ref}}$. This reference point is taken to be the gas phase optimised conformational minimum of the ion. The Taylor expansion about the reference point may be written as follows

$$\Delta E^{intra}(\theta) \approx \Delta E^{intra}(\theta_{ref}) + \left[\frac{\partial E^{intra}}{\partial \theta}\right]^{T}_{\theta_{ref}} (\theta - \theta_{ref}) + \frac{1}{2}(\theta - \theta_{ref})^{T}\left[\frac{\partial^2 E^{intra}}{\partial \theta^2}\right]_{\theta_{ref}} (\theta - \theta_{ref}) \qquad \textbf{(23)}$$

where ΔE^{intra} is the intramolecular energy of a given conformation relative to the input gas phase optimised geometry. This reduces the computational costs associated with estimating the intramolecular energy differences between different conformations. The other problem that needs to be addressed is the conformational dependence of the multipole model used to describe the intermolecular interactions between the ions in a crystal and the computational cost of re-calculating the molecular charge density at each outer optimisation step. It has been shown[122] that this computational cost can be significantly reduced by rotating the multipole moments with the local environment. This approximation is valid because the LAM only operates within a small region of the molecular conformational space and if the change in $\boldsymbol{\theta^f}$ is greater than a limit of ± 4°, CrystalOptimizer re-constructs the LAM by performing a constrained geometry optimisation of the molecular conformation (with torsions fixed to last point). It then uses the resulting gas phase minimum conformation to calculate a new molecular charge density. CrystalOptimizer uses a database to store the intramolecular energies and multipole moments for each LAM. When refining multiple computer generated crystal structures of the same system, this can lead to significant savings in time.

2.5.4 Estimating the Helmholtz free energies of organic crystals

So far we have only discussed the stabilities of organic crystals in terms of the static

0 K lattice energies estimated by DMACRYS. However, these idealised perfect crystal simulations neglect the effects of temperature and a more rigorous treatment would require the Gibbs free energy (*G*) surface rather than the lattice energy surface commonly used in crystal structure prediction studies. Thus the static 0 K lattice energies used in crystal structure prediction neglect the entropic and zero-point energies of crystals. The effects of temperature are absorbed into the potential parameters of the FIT and W99 force fields, which have been fitted to experimental crystal structures determined at various temperatures. It is possible to estimate the effects of dynamic motions of the rigid body molecules within DMACRYS. This is done by calculating the phonon frequencies of the lattice energy minimum under the harmonic approximation. Because the molecules are kept rigid during lattice energy minimisation with DMACRYS, only the intermolecular lattice vibrations are treated and so the only motions sampled are those arising from translation or libration of the rigid body entities. The starting point[123] for calculating the phonon frequencies is the dynamical equation[123]

$$D^{ij}_{\tau\tau'}(\boldsymbol{k}) = \left(\boldsymbol{M}_{\tau,i}\boldsymbol{M}_{\tau',j}\right)^{-\frac{1}{2}} \sum_{l'} \Phi^{0l'}_{\tau\tau'}(ij) \exp(i\boldsymbol{k}\cdot\boldsymbol{r}^{l'}) \qquad \textbf{(24)}$$

where $\boldsymbol{M}$ are the molecular masses (for translational motions) or moments of inertia (for rotations) and $\Phi^{0l'}_{\tau\tau'}(ij)$ is the matrix of second derivatives with respect to the displacements of molecule *i* (in the reference unit cell) and molecule *j* (in unit cell *l'*) along τ and τ' respectively. The phonon frequencies are calculated by using the above dynamical equation to solve the following eigenvalue problem[123]

$$\left| D^{ij}_{\tau\tau'}(\boldsymbol{k}) - \omega^2(\boldsymbol{k})\delta_{\tau\tau'}\delta_{ij} \right| = 0 \qquad \textbf{(25)}$$

where $\omega(\boldsymbol{k})$ is the phonon frequency. DMACRYS calculates the phonon frequencies for $\boldsymbol{k} = 0$ and this reduces the mathematical complexity of evaluating the phonon frequencies considerably since the exponential term in equation 24 vanishes at $\boldsymbol{k} = 0$. Extension to $\boldsymbol{k} \neq 0$ would require a non-trivial adaptation of the Ewald sums for the electrostatic interactions to take into account the $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ term in equation 24. The phonon frequencies provide access to the vibrational energies of crystals. The starting

point for evaluating the vibrational contributions towards the total crystal energy is the vibrational partition function of the harmonic oscillator. For a crystal, composed of N unit cells, each consisting of n atoms, the vibrational partition function may be defined as[123]

$$\ln Z_{vib} = -\frac{1}{2}\sum_{i=1}^{3nN}\frac{\hbar\omega_i}{kT} - \sum_{i=1}^{3nN}\ln\left[1-\exp\left(-\frac{\hbar\omega_i}{kT}\right)\right] \quad \textbf{(26)}$$

and the link to the free energy is given by the expression[123]

$$G = \Phi - kT\ln Z + pV = \Phi + \frac{1}{2N_k}\sum_{\boldsymbol{k}j}\hbar\omega_j(\boldsymbol{k}) + \frac{kT}{N_k}\sum_{\boldsymbol{k}j}\ln\left[1-\exp\left(-\frac{\hbar\omega_j(\boldsymbol{k})}{kT}\right)\right] + pV \quad \textbf{(27)}$$

where Φ is the rigid body lattice energy and N_k is the number of **k**-points. It is much more convenient however to work with the Helmholtz free energy surface since under normal laboratory conditions, the difference in crystal densities between polymorphs is small and the $p\Delta V$ contribution towards the free energy differences are negligible (see section 1.2.3). Thus DMACRYS computes the Helmholtz free energy as[123]

$$A = \Phi - kT\ln Z = \Phi + \frac{1}{2N_k}\sum_{\boldsymbol{k}j}\hbar\omega_j(\boldsymbol{k}) + \frac{kT}{N_k}\sum_{\boldsymbol{k}j}\ln\left[1-\exp\left(-\frac{\hbar\omega_j(\boldsymbol{k})}{kT}\right)\right] \quad \textbf{(28)}$$

The free energy differences between predicted structures that are estimated in this way are sensitive[123] to the model intermolecular potential parameters used.

2.6 Previous work on the crystal structure prediction of organic salts

Recent years have seen increasing interest in the *ab initio* prediction of the crystal structures of multi-component solid forms, with most of the work in the literature devoted to cocrystal[122,124-126] solid forms. The crystal structures of hydrates[80] and solvates[79] have also received some attention. By contrast the *ab initio* prediction of the crystal structures of organic salts has received very little attention and this provided the

motivation for the research presented in this thesis. The greater body of work on the crystal structure prediction of cocrystals, is a result of the recent interest[17,127-130] such solid forms have received in the crystal engineering community. The tendency for salts to crystallise as hydrates and the structural ambiguities that can arise from proton disorder suggests that the blind prediction of their crystal structures will be challenging. However, there have been some early efforts that are worth mentioning here. Karamertzanis *et al*[131] have performed the crystal structure prediction of a set of diastereomeric salts and have tested the utility of the predicted lattice energies in explaining the separation efficiency without prior experimental input. The authors used the static 0 K lattice energies to rank the predicted structures using the FIT force field and the balance between the inter- and intra-molecular energies of the flexible ions were estimated using the DMAREL[132]-DMAflex[121] suite of programs. The authors were successful in obtaining qualitative agreement between the computed and experimental relative energies of the structures for the polymorphic (*R*)-1-phenylethylammonium-(*S*&*R*)-2-phenylbutyrate system. However, the crystal energy landscape of (*R*)-1-phenylethylammonium-(*S*)-2-phenylbutyrate suggested a relative lattice energy of 40 kJ mol^{-1} between the *RS* salt and most stable *RR* polymorph. This is a significant overestimate of the actual energy difference between these diastereomeric salts as estimated from the solubility differences[131]. Antionidis[133] later performed an extensive polymorph screen for the *RS* salt in the hope of finding more stable experimental structures that match one of the predicted low energy structures. The experimental screen indicated that other metastable forms of the salt do exist but these were not long lived so as to allow their characterisation via diffraction methods. Leusen[134] has also applied computational crystal structure prediction to the problem of racemate resolution via preferential crystallisation of diastereomeric salts. As part of this work, Leusen studied two diastereomeric salts consisting of a chlorine-substituted cyclic phosphoric acid and the two enantiomers of ephedrine.

McArdle *et al*[135] have attempted the crystal structure prediction of chloride and bromide salts using the MOLPAK-WMIN[110,136] suite of programs. The lattice energy of the hypothetical structures were calculated using the atom-atom potential parameters implemented in WMIN[136] in conjunction with a semi-empirically derived Cl potential[137]. Given the origin of these intermolecular potentials, it was not surprising that McArdle was predicting the known structures of *o*-toludinium chloride and *o*-

toludinium bromide at least 10 kJ mol^{-1} higher in energy than the global minimum from the search. More recent work by Van de Streek *et al*[138] has used a proprietary dispersion-corrected density functional theory method[139] as implemented in GRACE[140] to model the polymorphs of pyridinium chloride[141-143]. Despite the fact that the dispersion-correction was parameterised against the crystal structures of neutral organic compounds, the method was successful in predicting the known phases of the salt as the three lowest energy structures in the crystal energy landscape. However, this success was not replicated in phenylethylammonium lactate[144], whose only known crystal structure was predicted as rank 5 in the crystal energy landscape of the salt. The authors found this result 'disappointing' but it suggests that the parameterisation of the dispersion-correction is inadequate for salts with many competing intermolecular interactions in the crystal (see Chapter 7, section 7.5).

3 Using the Cambridge Structural Database as a tool to understand the structures of organic salts and cocrystals

3.1 Introduction

The Cambridge Structural Database[29] (CSD) is widely acknowledged as being the most comprehensive depository of the crystal structures of small organic/organo-metallic systems. The CSD is managed by the Cambridge Crystallographic Data Centre (CCDC). Compilation of the CSD began in 1965,[29] with the focus of developing a central database for storing all relevant crystallographic information for each published crystal structure. At a basic level, this information includes the unit cell dimensions, atomic co-ordinates (where available) and relevant bibliographic information. As of May 2010, the CSD contained a total of 512,881 crystal structures. The growth of the CSD has benefited from the advances made in the data processing of the results obtained from diffraction experiments. This in itself has only been possible because of the advances made in computing power over the past quarter of a century. The CSD records full 3D structures obtained from X-ray/neutron diffraction experiments and the database contains structures solved from both single crystal and powder samples. Sometimes full 3D atomic co-ordinates are not obtained from the diffraction experiment, and in such cases the CCDC only archives the unit cell parameters and any other experimental information available on the structure.

All releases of the CSD include the visualisation package *Mercury*[114,145], which allows standard as well as advanced functionality for viewing the crystal structures obtained from a typical search. The search itself is performed using the program *ConQuest*, which implements text, sketch and numerical search fields for finding the requested information. The availability of many crystal structures in the CSD, consisting of molecules with the same functional groups and hence intermolecular contacts, has made it possible to derive simple isotropic atom-atom potentials[146,147] for modelling the crystal structures of organic compounds. Only a dozen well determined structures are typically used in the fitting of these atom-atom potentials. Whilst this is an important application of the CSD, theoreticians represent a very small community of the people who use the database and they are certainly outnumbered by crystal engineers and other solid state chemists who routinely use the database in their work.

In addition to the obvious role the CSD plays in the storage of crystal structures, the wealth of structural information contained in the database makes it amenable to statistical analyses. There are some general caveats associated with searching the CSD and any meaningful statistical survey of the database must be performed in light of these caveats. The first caveat relates to the quality of some of the older structures in the database, which if subjected to the stringent tests[148] now routinely used to validate *crystallographic information files* (CIFs[149]), could legitimately be described as poorly determined or even erroneous[150]. Structural errors are likely among these older structures because they were keyboarded[150] twice: once during publication in a journal and again when recorded in the CSD. If the CSD survey draws upon structural parameters such as torsion angles or intermolecular distances, filters should be used to eliminate poorly determined structures, otherwise there is the risk of sampling unrealistic values for the search parameter. A second caveat is that the search fragment defined in ConQuest need not always yield structures that are consistent with the sketch. For example, in metal containing structures it is known[151,152] that there can sometimes be inconsistencies between the charges reported by ConQuest and those implied by the co-ordination number of the metal in the crystal structure. If one is interested in comparing the occurrence frequencies of two types of structures, these inconsistencies can introduce errors in the frequencies. Other more specific caveats will be mentioned under the appropriate section of this chapter. The aim of this chapter is to summarise the results of published CSD surveys that highlight key trends in the structures and properties of organic salts and cocrystals. Suitable case studies of how the information from such surveys has been used to crystallise novel solid forms will also be presented. Where applicable, the results obtained from inspection of other databases will be discussed and new results of CSD surveys on particular problems related to organic salt crystal structures will be presented.

3.2 Applications of the CSD to the design of cocrystals

In the past few years, a great deal of attention has been given to the "design" of multi-component solid forms using the principles of crystal engineering[153], with the majority of publications in the field focussing on cocrystals[14,17,129,154]. Here the word "design" refers to the synthesis of solid forms with desirable structures and ultimately properties.

Zaworotko[129] has noted that the 'analysis of existing crystal structures represents the first step in a crystal engineering experiment', and the CSD has been extensively used[155-157] in this endeavour. Central to the application of the CSD to the design of novel cocrystals is the idea that if we understand the crystal packing of a large set of known structures, we can use this information to help target novel structures. This is possible when the coformers used in the crystallisation experiments have the same functional groups as the coformers sampled in the CSD survey. The procedure assumes that crystallisation is driven primarily by the strength and directionality of hydrogen bonds but says nothing about the effects of differences in the chemical properties between the coformers sampled in the CSD survey and those actually used in the laboratory.

There has been a great deal[127,158,159] of emphasis placed on the importance of hydrogen bond donor and acceptor groups in driving molecular complementarity, despite the many subtle intermolecular forces at work in producing the crystal structure and the inadequacy of hydrogen bonds to describe the origin or impact of such forces. Most cocrystal design strategies are based on the idea that hydrogen bonds form in a hierarchical[22,23,160] fashion (see section 1.2.2). There is a great deal of empirical evidence in the CSD to support this and one example is the high occurrence[161,162] frequency of the $CO_2H\cdots N_{arom}$ heterosynthon in cocrystals formed from carboxylic acids and simple pyridines. The first reported example of supramolecular synthon polymorphism[163] in cocrystals has also shown that the more stable form displays the expected supramolecular heterosynthon based on the hierarchy of supramolecular synthons, while the metastable form crystallises with homosynthons.

Fabian[164] has attempted to address the factors that affect molecular complementarity in cocrystals. He did this by analysing the structures of 974 cocrystals retrieved from the CSD. A set of 131 molecular descriptors were used for each molecule in a cocrystal. These included hydrogen bond donor and acceptor counts, descriptors for shape, size and molecular electrostatic properties. A statistical analysis of the data was subsequently performed to identify the most strongly correlated descriptors in molecules of each cocrystal. The results showed that the majority of cocrystals are formed from molecules of similar shape and polarities. By contrast, no significant correlation was found between the number of hydrogen bond donor and acceptor groups. This is not to say that hydrogen bonds are unimportant in the design of

novel cocrystals, but rather that we need to understand the full range of factors that drive molecular complementarity if we are to be in a position to influence the resulting structure of the cocrystal.

Following a careful search of the CSD, it has been suggested by Steed *et al*[157] that molecules which crystallise with Z′>1 show a 'markedly stronger tendency to form cocrystals' than those that crystallise with Z′=1. The search of the CSD revealed that only 5 % of bioactive molecules with Z′=1 form cocrystals while the number goes up to 7 % when bioactive molecules with Z′>1 are considered. The results of the statistical survey were put to test by the authors and proved valid. This was done by synthesising a novel cocrystal from a bioactive molecule with Z′>1 and also by crystallising a previously unknown Z′>1 structure from a molecule known to crystallise only as a cocrystal. The authors suggested that although the reasons that lead to Z′>1 crystallisation are complicated[165-167], one of the possible causes is non-self complementarity of a system with regards to molecular shape and/or functionality. If this factor is indeed influential, it follows that molecules with Z′>1 structures are less likely to pack efficiently with themselves and hence will be more effective cocrystallising agents than molecules (Z′=1) that have a greater degree of self-complementarity in the solid state. This idea of non-self complementarity has also been used to suggest[168,169] that polymorphic compounds are better sources for cocrystals when compared to non-polymorphic compounds.

The utility of the CSD for aiding experimental screens that target novel cocrystals is further illustrated by the work of Sadiq *et al*[170], which suggests that Hammett[171] constants can be used to guide the synthesis of cocrystals with specific hydrogen bond motifs. In the Hammett equation

$$\log(K) = \rho\sigma + \log(K_0) \qquad \textbf{(29)}$$

both the sign and numerical value of the Hammett constant, σ, is a measure of 'the extent of electron donating (σ negative) and withdrawing (σ positive) capability of a substituent on a phenyl ring, relative to benzoic acid as a reference state.' In equation 29, K denotes the equilibrium constant for the reaction involving the substituted benzoic acid derivative and K_0 refers to the equilibrium constant for the reaction

involving benzoic acid. The symbol ρ denotes the reaction constant characteristic of the substituent studied. The Hammettt equation has been widely applied in studying the effects of substituents on the kinetics of reactions involving the making and breaking of covalent bonds but was not intended to be used in supramolecular "reactions" which do not involve the making and breaking of such bonds. However in the work of Sadiq *et al*, no attempts were made to try and obtain kinetic data and instead the σ values were used to match acid molecules that were good hydrogen bond donors with those that were good hydrogen bond acceptors. A CSD survey was performed to see if there is a relationship between Hammett constants for particular functional groups attached to benzoic acid and the success rate of forming a cocrystal displaying the carboxylic dimer hydrogen bond motif. From the CSD survey, the authors learned that cocrystals displaying the desired carboxylic acid dimer motif were frequently formed from benzoic acid derivatives where one coformer had an electron donating (typically amine) group while the other had an electron withdrawing (typically nitro) group. Following this finding, they attempted to form 1:1 cocrystals of substituted benzoic acid derivatives and the coformers were chosen on the basis of the Hammett constants of substituents being comparable in value but opposite in sign. The authors successfully synthesised two novel cocrystals displaying the expected carboxylic acid dimer motif. Because of the influence of steric effects, the Hammett equation cannot be applied for compounds with substituents *ortho* to the carboxylic acid group. Indeed, the only attempted cocrystallisation experiment, involving a coformer with an *ortho* substituted group, led to a salt displaying a supramolecular homosynthon between molecules of the cation. The anion was found to be involved in discrete hydrogen bonds with the cation via a carboxylate-ammonium interaction. The work of Sadiq *et al*[170] has shown that the application of the CSD need not always start from a detailed statistical analysis and end with the framing of experimentally testable hypotheses. In fact the process can also be done in reverse, and the CSD can be used to confirm or reject a hypothesis based on empirical data.

3.3 Known CSD trends in the structures of molecular organic salts

3.3.1 Overview

While a great deal of work has appeared over the past decade detailing the results of CSD surveys on cocrystals, there have been far fewer papers reporting the results of CSD surveys on salts and even fewer reporting the results pertaining specifically to pharmaceutically acceptable salts[172]. Given that salt formation is considered one of the most widely[66,173] used methods for modifying the physicochemical properties (*e.g.* solubility, melting point, dissolution rate) of a drug molecule, this may be surprising, but reflects the current interest in cocrystal solid forms. This section provides a summary of published work that has used the CSD to quantify the frequencies of pharmaceutically acceptable counterions as well as the hydration behaviours of organic salts as a function of counterion. In section 3.4, the results of novel CSD surveys on organic salts are presented.

3.3.2 The occurrence frequencies of pharmaceutically acceptable ions

Haynes[152] has analysed the distribution of pharmaceutically acceptable cations and anions in the CSD using a list of 69 acids and 21 bases taken from the *Handbook of Pharmaceutical Salts*[64]. Because the CSD does not exclusively compile the crystal structures of pharmaceutically acceptable molecules or their salts and the variety of molecules in the database is far greater, one must be careful of the conclusions drawn from such a search. The CSD does have the facility to restrict searches to the crystal structures of bioactive molecules and this is done by including the text string 'BIOACTIVE' as part of the search. In the work of Haynes, it was found that only 19 % of molecules classed as bioactive existed as salts containing a positively charged N atom. Haynes found that of the list of acids and bases investigated for their frequency of salt formation in the CSD, the most common anion was the chloride (Cl^-) accounting for 46 % of the total number of hits (this includes hits for mono- and di-ions). If we just consider the 15 anions listed in Table 3.1, the pie chart of Figure 3.1 shows that the Cl^- anion accounts for 54 % of the total number of hits. The second most frequently occurring anion in organic salts of the CSD is the bromide (26 %). Of the cations of pharmaceutically acceptable bases searched for, the ammonium (NH_4^+) cation was the

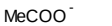
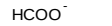
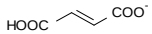
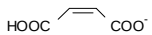
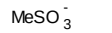
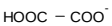
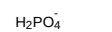
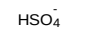
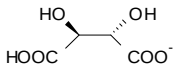
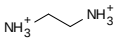
Ion	Chemical sketch	Number of unique structures	Number of hydrates	% Hydrates	Number of polymorphs	% polymorphs
Acetate		61	27	44	5	8
Bromide		1403	346	25	31	2
Camsylate		29	6	21	0	0
Chloride		2874	894	31	77	3
Formate		30	11	37	2	7
Fumarate		36	5	14	0	0
Maleate		78	8	10	2	3
Mesylate		63	20	32	5	8
Nitrate		290	559	20	14	5
Oxalate		61	18	30	4	7
Phosphate		97	17	18	6	6
Sulfate		70	17	24	4	6
Tartrate (+L)		86	32	37	4	5
Thiocyanate		73	13	18	0	0
Tosylate		118	27	23	9	8
Ammonium		265	85	32	10	4
Ethylenediamine		83	30	36	5	6

Table 3.1: Tabulated occurrence frequencies of selected pharmaceutically acceptable anions (above grey line) as well as the two most frequently occurring cations (below grey line) found in pharmaceutically acceptable salts of the CSD. The data was reproduced from the CSD survey of Haynes[152].

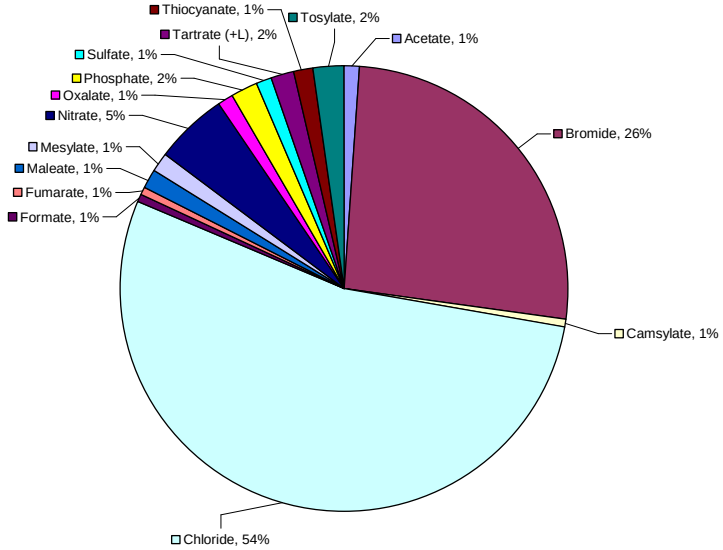


Figure 3.1: The % occurrence of pharmaceutically acceptable anions in organic salts of the CSD as found from the survey of Haynes[152]. The pie chart shows the number of each type of anion listed in Table 3.1 expressed as a percentage of the total number of hits for these anions.

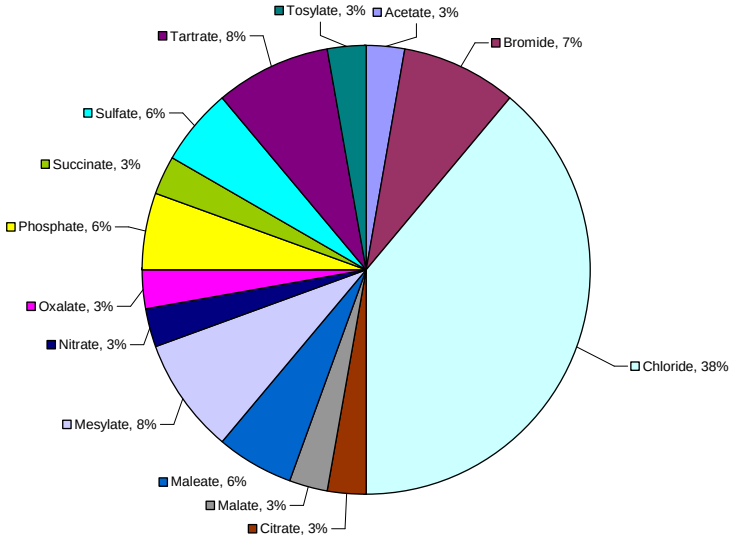


Figure 3.2: The % occurrence of anions in FDA approved pharmaceutical salts of the Orange Book. The above pie chart is based on the survey of Saal[172] and reflects the contents of the Orange Book for the period 2002-2006.

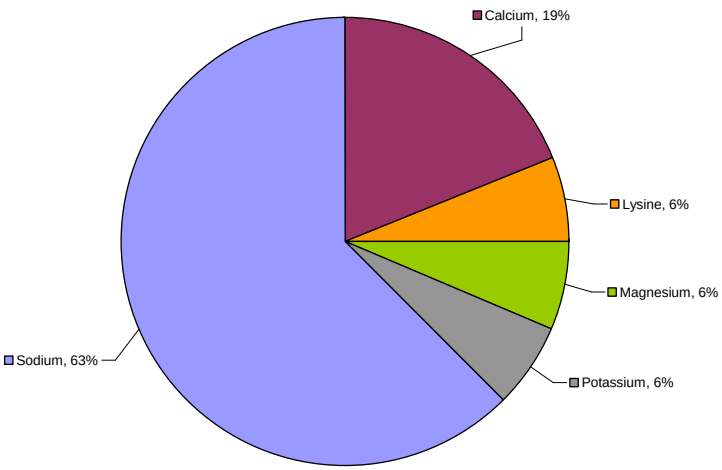


Figure 3.3: The % occurrence of cations in FDA approved pharmaceutical salts of the Orange Book. The above pie chart is based on the survey of Saal[172] and reflects the contents of the Orange Book for the period 2002-2006.

most prevalent.

Saal *et al*[172] looked at the distribution of pharmaceutically acceptable counterions in the Orange Book Database[174] published by the U.S. Food and Drug Administration (FDA) agency. Unlike the CSD, the list of salts stored in the Orange Book contains only those that are classed as pharmaceutically acceptable. For a particular entry, other differences include the fact that the Orange Book does not provide the detailed structural information found in the CSD. The aim of the survey by Saal *et al*[172] was to obtain a reliable estimate of the proportion of pharmaceutically acceptable salts belonging to a particular class of counterion and contrast the results with those based on the CSD survey of Haynes[152]. Distribution trends over time were also analysed. Entries in the Orange Book prior to 1981 do not have a date of approval for the drug and so analyses of changes in the occurrence frequencies of certain counterions were limited to the period 1982-2006. From the survey, the authors found that the most prevalent anion in FDA approved pharmaceutical salts during the period 2002-2006 was the chloride anion (Figure 3.2) accounting for 38 % of the total drugs approved. The chloride anion was also the most frequently occurring anion in all four of the five year intervals from 1982-2001 and has an overall occurrence frequency of 54 %. The investigation of trends in the anion distributions over time revealed preferences for specific anions at

certain periods in history. For example, nitrates represented 8 % of approved salts during the period 1982-1986 although the average incidence of nitrates over the entire time period studied was only 2 %. Similarly, the occurrence frequency of acetate anions peaked at 13 % during the period 1987-1991 despite an overall occurrence frequency of 3 %. For the distribution of cations found in salts of APIs (Figure 3.3) limited to the period 2002-2006, the two most prevalent cations were Na^+, accounting for 63 % of the total, and Ca^{2+} accounting for 19 % of the total.

3.3.3 Trends in the hydration behaviour of N^+-H containing salts

The trends in the hydration behaviour of N^+-H containing salts as a function of pharmaceutically acceptable anion have been summarised by Haynes[73]. The anion types surveyed were based on acids taken from the *Handbook of Pharmaceutical Salts*[64] and can be summarised according to the groups: carboxylates and carbonates (CO_2^- /CO_3^{2-}), halides (X^-, X=halogen), sulfates and sulfonates (SO_4^{2-}/RSO_3^-), phosphates ($H_2PO_4^-$/HPO_4^{2-}/PO_4^{3-}), nitrates (NO_3^-) and thiocyanate (SCN^-). The investigation of trends in the hydration behaviour of these anion groups was limited to salts where the cation is formed from pyridine or an amine. The amine was either cyclic or non-cyclic (1°, 2° and 3° substitution at nitrogen). Figure 3.4 provides a plot of the % hydration in the above anion groups as a function of the cation group surveyed. The hydration behaviour of specific halide anions (not reported in Figure 3.4) was found to be size dependent, with the general trend of increasing hydrate formation with decreasing ionic radius of the halide anion. The F^- anion has the smallest ionic radius of all the elements in Group 17 and because of its greater polarising effect when compared to other halide anions has an increased likelihood of incorporating water molecules into the crystal lattice. Halide salts under all cation groups have % hydration values that are statistically significant ($\geq$ 5 %) and this peaks for the pyridinium cation at a value of 34.8 %. Of the other anion groups investigated by Haynes, phosphates and sulfates were found to have the highest % hydration under all cation groups with the exception of pyridinium. With the pyridinium cation, Haynes found that phosphates and halides have the highest % hydration, with sulfates coming in a close third. The high % hydration in phosphates and sulfates is in agreement with the observation by Lourdes *et al*[175] that an increase in the number of polar groups generally leads to an increase in the

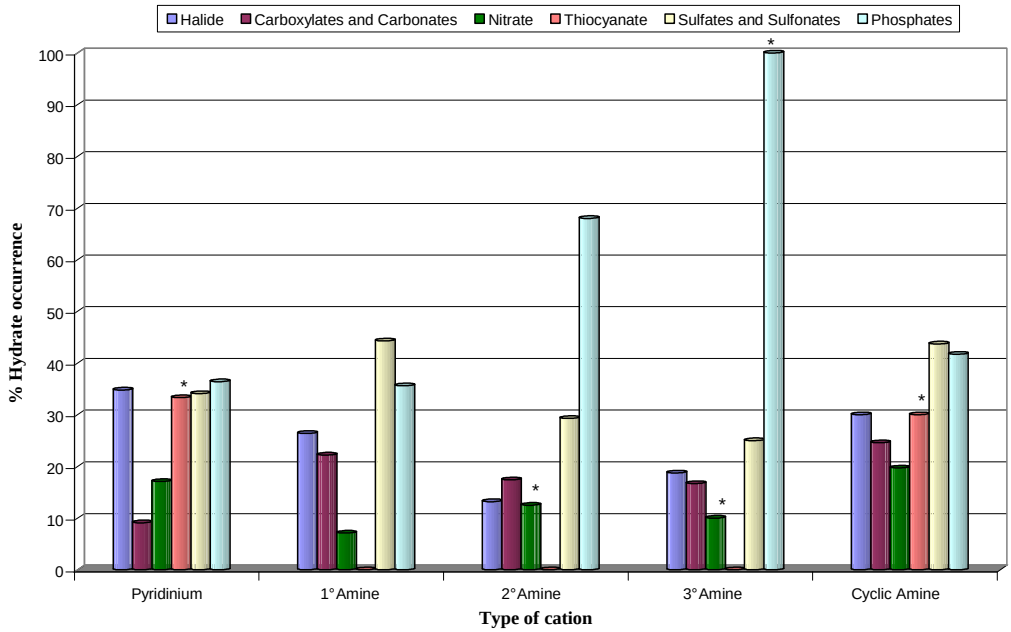


Figure 3.4: Percentage hydrate occurrence in NH^+ containing salts of the CSD that contain the anion groups: halide, carboxylates/carbonates, nitrate, thiocyanate, sulfates/sulfonates and phosphates. The above results are based on the survey of Haynes[73]. * For these categories of salt structures, the quoted % hydrate occurrence is based on a sample population of 10 crystal structures or less.

frequency of hydration. However, the value of 100 % hydrate occurrence for phosphates under the 3° amine category must be treated with caution as there were only 6 structures found from the CSD survey and this is not a large enough sample population to allow meaningful statistics to be obtained. Other cases where the % hydration values quoted by Haynes were based on a sample population of 10 crystal structures or less are indicated in Figure 3.4. Haynes observed that only 9 % of pyridinium carboxylate salts were hydrates and this was second lowest in % hydration to nitrate salts containing a 1° amine cation. Moreover, the reduced frequency of hydration in carboxylates was found to be limited to salts where the cation is specifically pyridinium, as indicated by the statistically significant higher % hydration found in cations based on primary, secondary and tertiary amines. Carboxylate salts with cation groups based on an aromatic amine (excluding $C_5H_6N^+$) or a general cyclic amine were found to have percentage hydration values that were more than a factor of two larger than those observed for pyridinium carboxylate salts.

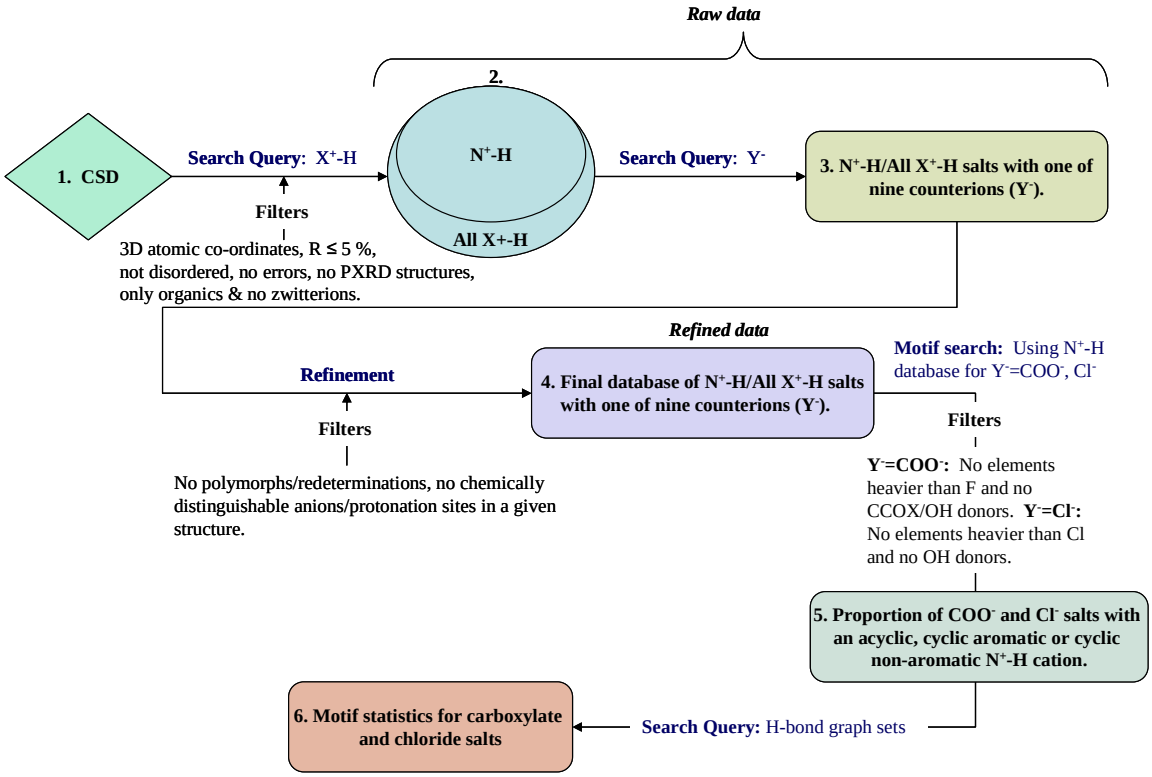
3.4 Survey of the CSD: Crystal structures of organic salts formed from acid-base proton transfer

3.4.1 Aim

The preceding sections of this chapter provide a summary of CSD data gathered from the literature, on some trends in the structures and properties of organic salts and cocrystals. In this section, the database will be used to answer two specific questions that relate to the crystal structures of organic salts. Initially, the CSD will be used to determine the most common proton acceptor in the crystal structures of organic salts formed from acid-base proton transfer. The database will then be used to investigate the types of hydrogen bond motifs adopted by carboxylate and chloride salts. These novel CSD surveys are motivated by the lack of published data on these topics.

3.4.2 Method

All Cambridge Structural Database surveys were performed with V5.30 of the CSD using the November 2008 update as implemented in Conquest V1.11. Searches were performed to elucidate the occurrence frequencies of certain anion groups, Y^-, as a function of the identity of the cation group, X^+-H. The following filters were used in the CSD searches: 3D atomic co-ordinates determined, $R \leq 5$ %, not disordered, no errors, no PXRD structures and only organics. Zwitterions were excluded by flagging all structures found by ConQuest, which contain an intramolecular contact between the charged groups X^+-H and Y^-, separated by 1-999 bonds and a distance in the range 1-40 Å. Initially, the CSD was searched for all structures containing the X^+-H (X=any element) group. This limited the survey to salt structures formed from acid-base proton transfer. The search was then repeated for heteroatoms from the *p*-block of the Periodic Table that are directly bonded to an acidic proton and bear a positive charge. The structures from the searches involving the X^+-H (X=any element) and N^+-H groups were saved as local databases and the occurrence frequencies of the following counterions investigated: COO^-, Cl^-, F^-, Br^-, NO_3^-, N^-, $H_2PO_4^-$, RO^- (excluding COO^-) and HSO_4^-. Two categories of search results are reported. The "raw data" contains the number of CSD entries with a particular anion irrespective of whether the structures contain polymorphs or more than one chemically distinguishable anion or protonation



Scheme 3.1: Summary flow chart of the method used in the CSD searches described in this section.

site (X_1^+-H and X_2^+-H, where $X_1 \neq X_2$). The "refined data" refers to structures with only one type of anion and the count under each anion category contains no polymorphs or re-determinations. For polymorphs and re-determinations reported under the same *refcode family*, the "best" crystal structure was kept, as judged by the entry with the lowest reported refinement residual, R_1. Structures that contain chemically distinguishable protonation sites were also eliminated from the list of refined data.

For crystal structures that contain the specific anion types $Y^- = COO^-$, Cl^- in the presence of the general NH^+ cation group, the results from the refined list of structures were saved as local databases and the structures analysed to elucidate the observed hydrogen bond motifs. All motif searches were performed using the Materials Module functionality in Mercury CSD 2.2[114,176]. Before carrying out the motif searches, filters were applied to both the carboxylate and chloride databases. For the database of carboxylate salts, the filters were to remove structures that contain elements heavier than F and those that contain the CCOX/OH donor groups. For chloride salts, the filters used were to remove structures that contain elements heavier than Cl and those that contain the OH hydrogen bond donor. In both cases, the first filter ensured the removal

of heavy element containing salt structures such as bridged metal/metalloid containing systems. Filters removing structures with the CCOX/OH groups ensured that structures with hydrogen bond donor groups that were competitive to the N^+-H donor of interest were not sampled including solvent containing structures where the solvent is either an alcohol or water molecule. Many chloride salts contained no co-ordinates for the hydrogen atoms and although Mercury CSD 2.2 can still calculate the graph set and associated motif based on the donor-acceptor geometries, these structures were removed prior to the motif search. Salt structures based on macromolecular cations such as those derived from porphyrins were also removed. The proportion of carboxylate and chloride salts that contain an acyclic, cyclic aromatic or cyclic non-aromatic cation were determined and the hydrogen bond motifs subsequently analysed using Mercury CSD 2.2[114,176]. A flow chart summarising the entire method described in this section is given in Scheme 3.1.

3.4.3 The most prevalent proton acceptor in the crystal structures of organic salts

A total of 7324 crystal structures were found in the search for salts containing the general X^+-H (X=any element) cation group. The N^+-H group was found in 7129 (97 %) structures of the total 7324. When the search was restricted to X^+-H containing cations where X= *p*-block element (excluding C), a total of 7265 structures were found and the N^+-H group was found to account for 98 % of this total. The results show that the nitrogen atom is the most common proton acceptor in the crystal structures of organic salts formed from acid-base proton transfer as stored in the CSD. The number of structures found under each anion category assuming the N^+-H or X^+-H cation groups are given in Table 3.2 and a bar chart comparing the search results is shown in Figure 3.5. The survey also confirmed that in order of decreasing prevalence, other commonly found cation groups (Figure 3.6) with statistically significant number of structures (≥20) in the CSD include O^+-H (98 structures) and P^+-H (28 structures). For salt structures bearing the O^+-H group, it was found that 62 % had the cation in the form of the hydroxonium ion (H_3O^+), 20 % in the form of a protonated carbonyl ($C=O^+$-H) and only 3 % with the proton directly bonded to an oxygen atom in an ether. There were no

	N^+-H		X^+-H	
	Raw	**Refined**	**Raw**	**Refined**
R-COO$^-$	1397	1306	1402	1313
Cl$^-$	1557	1406	1578	1425
F$^-$	38	30	39	31
Br$^-$	445	421	461	439
NO_3^-	180	149	184	152
N$^-$	205	189	212	196
$H_2PO_4^-$	155	130	158	130
R-O$^-$	302	289	308	293
HSO_4^-	46	37	47	41
Other	794	759	860	825

Table 3.2: Numerical results from the CSD search for salt structures with the group N^+-H or X^+-H (X=any element) in the ten anion categories considered. Both the raw and refined datasets are reported.

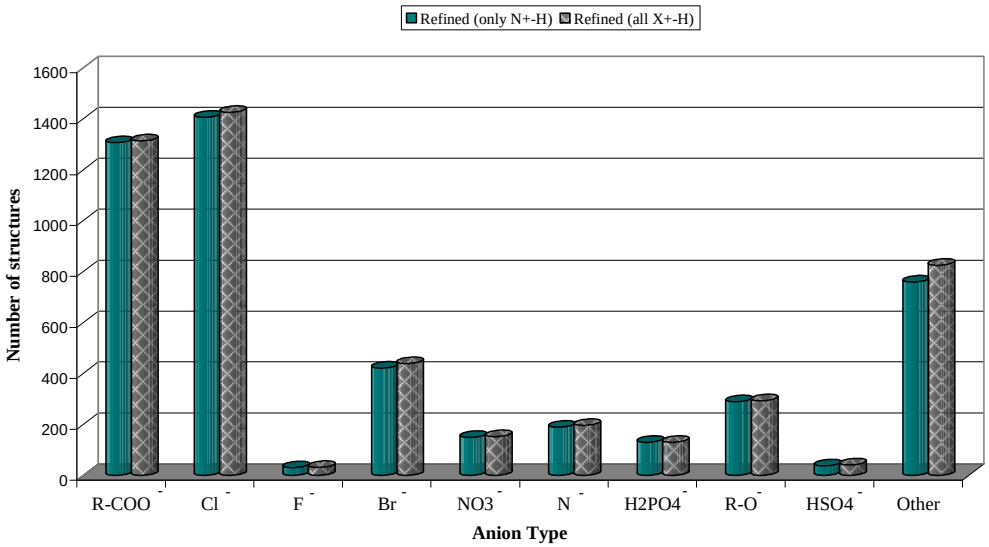


Figure 3.5: Comparison of the total number of N^+-H (teal) or X^+-H (hatched grey) containing salt structures with one of the following general anion categories: COO^-, Cl^-, F^-, Br^-, NO_3^-, N^-, $H_2PO_4^-$, RO^- (excluding COO^-), HSO_4^-. The figure shows that in each anion category, almost all the salt structures with a cation featuring the X^+-H (X=any element) group are of the type N^+-H.

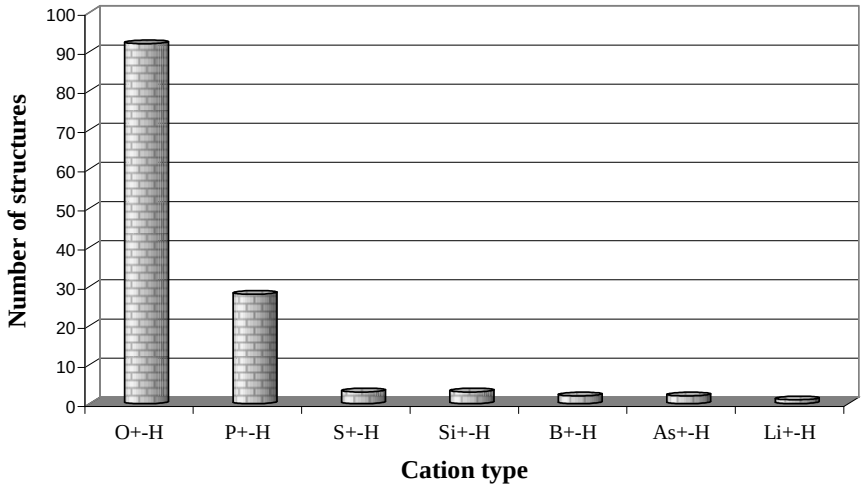


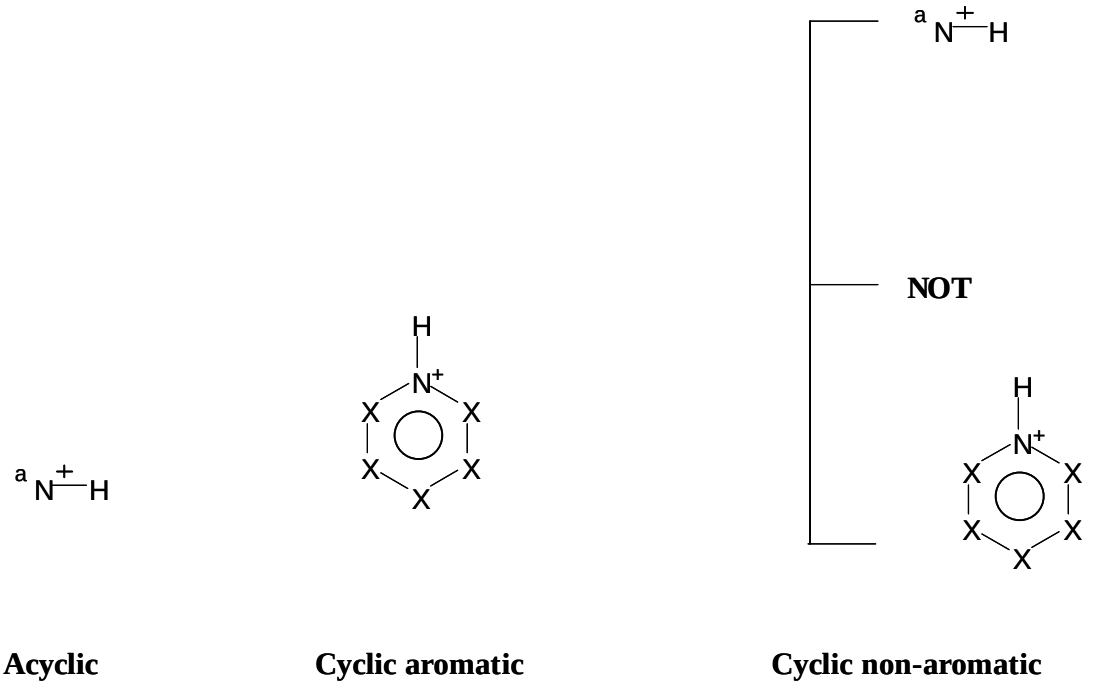
Figure 3.6: Illustration of the type and frequency of salt structures with proton acceptors other than the nitrogen atom as found in the local database of X^+-H type salt structures. Polymorphs and re-determinations were eliminated from the final tally plotted above. The figure shows that with the exception of the nitrogen atom, the two statistically significant (≥20 structures) cation groups found in the crystal structures of organic salts formed from acid-base proton transfer are O^+-H and P^+-H.

O^+-H containing salt structures derived from protonating an alcohol group.

The final tally (refined data) of salt structures for the general N^+-H cation group showed that the three most commonly occurring anions (excluding the category "other") are: Cl^- accounting for 1406 (36 %) structures, COO^- with 1306 (33 %) structures and Br^- accounting for 421 (11 %) structures. The finding that Cl^- is the most prevalent anion is in agreement with the work of Haynes[152] and Saal[172]. In the results reported here, specific carboxylate counterions (*i.e.* fumarate, maleate, acetate) were not searched for and the fact that the % occurrence of carboxylates is competitive with that of chlorides is a consequence of the search for the general group COO^-.

3.4.4 Observed hydrogen bond motifs in carboxylate and chloride salts

The carboxylate and chloride anions are more prevalent in organic salts of the CSD than any of the other counterions listed in Table 3.2. Because of the prevalence of these counterions, a search of the CSD was performed for the hydrogen bond motifs displayed by NH^+ containing organic salts where the anion is specifically carboxylate or chloride. The "refined" lists of chloride and carboxylate salts were used as the starting points for the motif searches and further filters (refer to section 3.4.2) were applied to remove structures with certain elements/functional groups that could potentially bias the



Scheme 3.2: Sketch of the ConQuest search fragments used to define the acyclic, cyclic aromatic or cyclic non-aromatic cation types in carboxylate and chloride salts of the CSD. [a] This superscript denotes that the N atom is *acyclic*.

motif searches. Following application of these filters for carboxylate salts, a final tally of 206 salt structures were obtained. Of the total 206 carboxylate salts, 134 (65 %) contained an acyclic cation, 39 (19 %) contained a cyclic (6-membered) aromatic cation and 33 (16 %) contained a cyclic non-aromatic cation. Scheme 3.2 shows the ConQuest search fragments used to define the cation types. For each of these cation types, Figure 3.7 shows a bar chart of the % number of carboxylate salts that display one of the five hydrogen bond motifs illustrated. In addition to the five hydrogen bond motifs considered in the survey (see Figure 3.7), other hydrogen bond networks were found in a number of structures. However, the motif search was restricted to the five hydrogen bond networks shown in Figure 3.7 because these motifs had statistically significant occurrence frequencies ($\geq$ 5 % of total structures for a given cation) across one or more of the cation types surveyed. The $R_4^2(8)$ hydrogen bond motif represents an exception in this regard, and was included so as to allow comparison with the motif search for chloride salts (Figure 3.8).

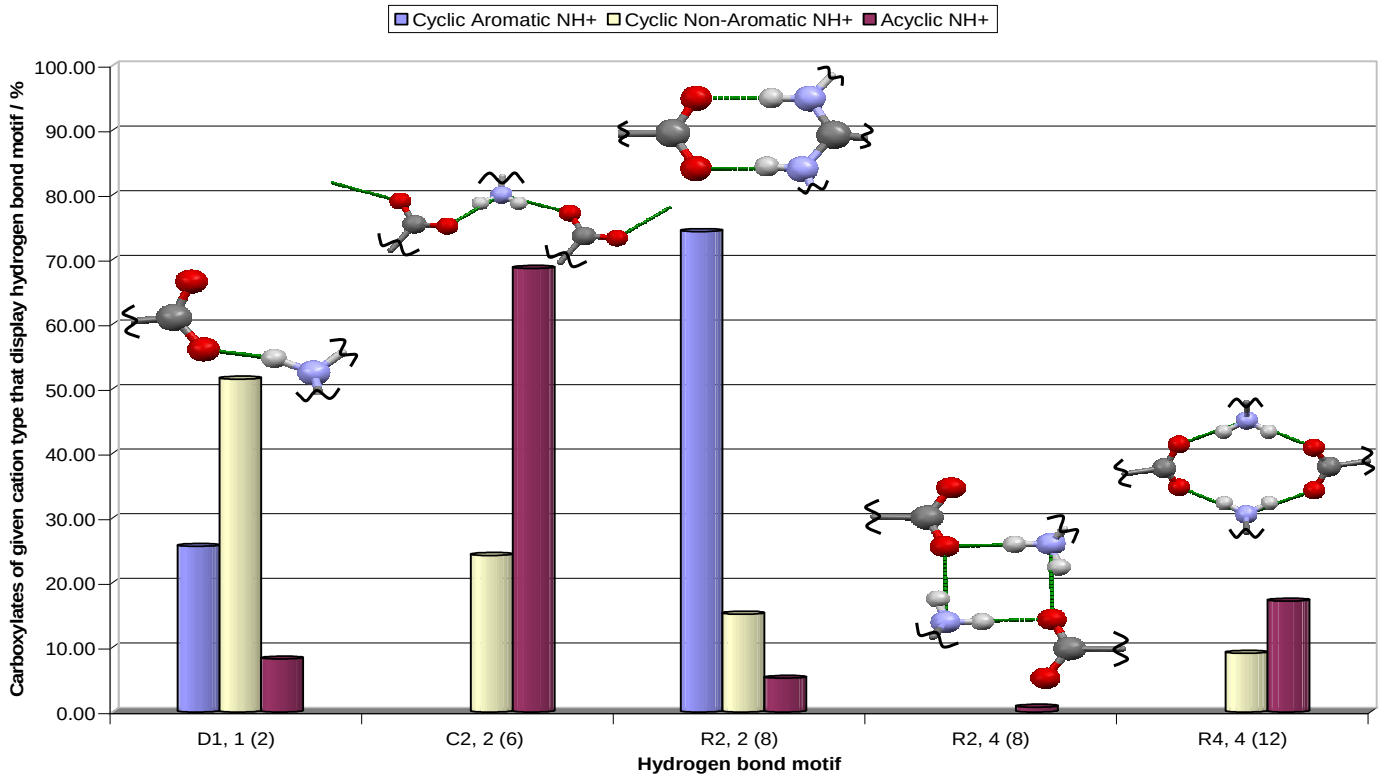


Figure 3.7: Hydrogen bond motifs found in carboxylate salts of the CSD. For a given cation type (acyclic, cyclic aromatic or cyclic non-aromatic), the chart shows the % number of structures that bear one of the five motifs illustrated. A curly line indicates a position where any substituent, R, can be attached and groups attached at these positions are not part of the hydrogen bond motif. The H-bond interactions in the motifs are illustrated by dashed green lines. In the figure, the notation[25] used for the graph sets is *G a, d (X)* where *G* is the designator, *a* and *d* are the number of hydrogen bond acceptors and donors and *X* is the degree of the hydrogen bond network. The D1,1(2) motif constitutes the simplest hydrogen bond interaction in organic salts that contain the COO^-/NH^+ ion pair. As such it can be found in all carboxylate salts irrespective of the cation category. However, only those structures that do not contain one of the other four hydrogen bond networks were listed as 'hits' under this motif.

According to Figure 3.7, the two most common hydrogen bond motifs found in carboxylate salts of the CSD are the heterodimer ring motif, $R_2^2(8)$, and the infinite ribbon motif, $C_2^2(6)$. The % occurrence of these motifs and by inference the likelihood of observing them in novel crystallisation experiments is strongly influenced by the type of cation found in the carboxylate salt structure. With cyclic aromatic cations, the $R_2^2(8)$ motif is the preferred hydrogen bond motif as judged by an occurrence rate of 74 %. This compares well with a recent survey of the CSD performed by Bis and Zaworotko[28], which showed that both the charge assisted $COO^-\cdots H^+NCNH_2$ and the neutral $COOH\cdots NCNH_2$ motifs have a combined occurrence probability of 77 % in crystallisation experiments involving a carboxylic acid and a 2-aminopyridine derivative. Allen *et al*[177] have also shown that in cocrystals formed from a carboxylic acid and a 2-aminopyridine derivative, there is a 76 % probability of observing the neutral $R_2^2(8)$ hydrogen bond motif in the resulting cocrystal. By contrast, the infinite ribbon motif, $C_2^2(6)$, can be found in carboxylate salts that contain either an acyclic or cyclic non-aromatic cation. However, the motif is almost three times as likely to be observed in salts that contain an acyclic cation (69 % occurrence rate) than is the case with those that contain a cyclic non-aromatic cation (24 % occurrence rate). The $C_2^2(6)$ hydrogen bond motif was not found in carboxylate salts that contain a cyclic aromatic cation and this can be explained when one realises that a requirement for the formation of an infinite ribbon motif is the presence of an NH_2^+ group, a feature that cyclic aromatic cations lack because of the sp^2 hybridisation at N. The majority (52 %) of carboxylate salts that contain a cyclic non-aromatic cation display the $D_1^1(2)$ motif.

A similar analysis of the CSD was performed to see the hydrogen bond motifs adopted by chloride salts. As for carboxylate salts, a motif was included in the survey only if it had an occurrence rate of at least 5 % under one or more of the cation groups - acyclic, cyclic aromatic or cyclic non-aromatic - of interest. Only three hydrogen bond motifs were found that satisfy this criterion and a bar chart of the % number of chloride salts that bear one of these motifs as a function of the cation type in the structure is given in Figure 3.8. The same search fragments (Scheme 3.2) used to define the cation types of carboxylate salts were also used for chloride salts and this allowed us to determine the number of chloride salts that fall under each of the three cation groups of

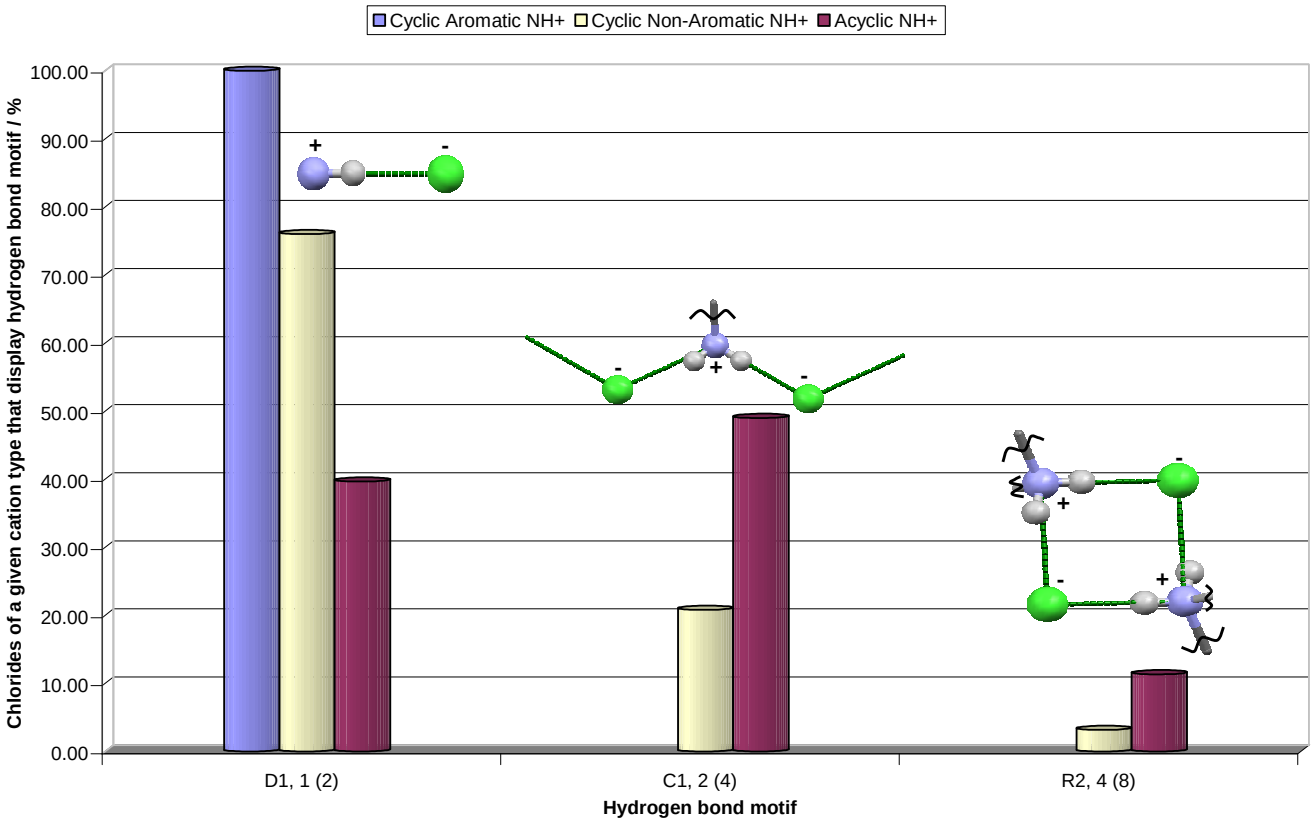


Figure 3.8: Hydrogen bond motifs found in chloride salts of the CSD. For a given cation type (acyclic, cyclic aromatic and cyclic non-aromatic), the chart shows the % number of structures that bear one of the three statistically prevalent hydrogen bond motifs illustrated. A curly line indicates a position where any substituent, R, can be attached, and groups attached at these positions are not part of the hydrogen bond motif. Graph sets are denoted *G a, d (n)* where *G* is the designator, *a* and *d* are the number of hydrogen bond acceptors and donors and *n* is the degree of the hydrogen bond network.

interest. There were 300 (50 %) chloride salts with an acyclic cation, 221 (39 %) with a cyclic non-aromatic cation and only 79 (13 %) chloride salts with a cyclic aromatic cation. For chloride salts with either a cyclic aromatic or cyclic non-aromatic cation, the most frequently occurring hydrogen bond motif was the discrete N^+-H···Cl interaction of graph set $D_1^1(2)$. This motif was found to have an occurrence rate of 100 % in chloride salts with a cyclic aromatic cation and this dropped to 76 % when salts with a cyclic non-aromatic cation were considered. For chloride salts with an acyclic cation, the most frequently occurring (48 % incidence) hydrogen bond motif was based on an infinite ribbon motif of graph set $C_2^1(4)$. This can be contrasted with the results for carboxylate salts (Figure 3.7) of the same cation type where the infinite ribbon motif also has the highest % occurrence. The only difference appears to be that in chloride salts, the $D_1^1(2)$ motif is more competitive with the infinite ribbon than is found for carboxylates. The $R_4^2(8)$ motif which occurs in less than 1 % of carboxylate salts (Figure 3.7) with an acyclic cation has an incidence rate of 11 % among chloride salts of the same cation group.

3.5 Conclusions

A survey of the literature has revealed that the chloride and carboxylate anions are the two most frequently used counterions in pharmaceutically acceptable salts stored in the Cambridge Structural Database (CSD). The literature also suggests that in general, chloride salts have a greater propensity for forming hydrates when compared to carboxylate salts. The finding that chlorides and carboxylates are the two most frequently used counterions in salts derived from organic bases is significant as it sets the direction for the types of systems studied in the remaining chapters of this thesis. Chapters 4 and 5 will present the results of work on selected carboxylate systems, while Chapter 6 will present the results of work on the computational modelling of chloride salts. The first salt system proposed as a challenge in the international blind tests of crystal structure prediction (Chapter 7) is also a carboxylate salt. Thus the CSD has served a vital role in quantifying the frequencies of salt counterions and has facilitated the choice of salts for validating the computational method of crystal structure prediction.

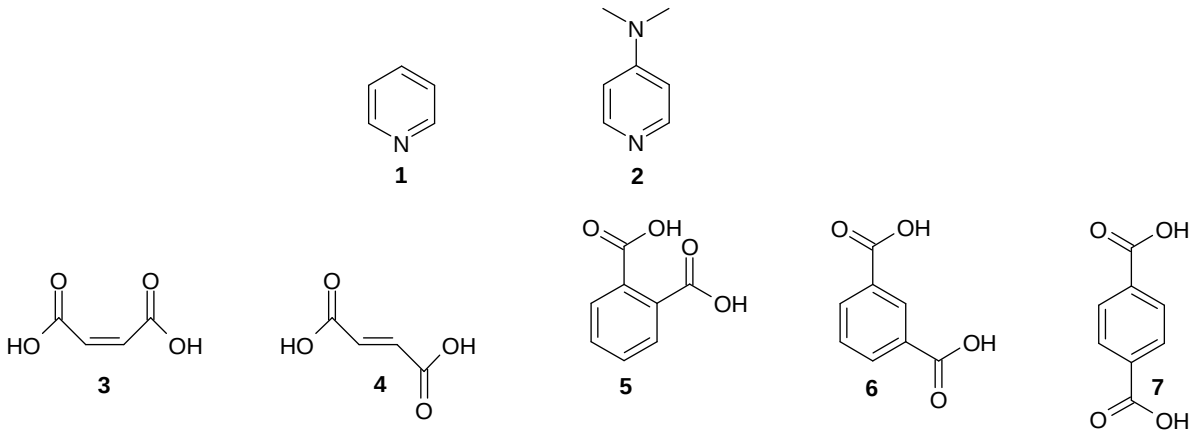
Novel CSD surveys were performed to quantify the most common proton acceptor (X^+-H) in organic salts and the most common hydrogen bond motifs displayed by chloride and carboxylate salts. Nitrogenous bases were by far the most common molecules used to crystallise organic salts and as such the N^+-H group was found in approximately 97 % of all salt structures surveyed. The surveys revealed clear preferences for certain hydrogen bond motifs when the anion crystallises in the presence of an acyclic, cyclic aromatic or a cyclic non-aromatic cation. For carboxylate salts with a cyclic aromatic cation, 74 % of structures adopt the $R_2^2(8)$ motif between the two ions. If the carboxylate salt consists of an acyclic cation, it is more likely to adopt the infinite ribbon motif, $C_2^2(6)$, with 69 % of structures with this cation type adopting this motif. The majority of chloride salts with a cyclic non-aromatic (76 %) cation and all chloride salts with a cyclic aromatic cation adopt the $D_1^1(2)$ motif. For chloride salts with an acyclic cation, 48 % of structures adopt the $C_2^1(4)$ infinite ribbon motif.

4 Salt or cocrystal? A new series of crystal structures formed from simple pyridines and carboxylic acids

4.1 Introduction

The rational[178] synthesis of cocrystal solid forms has received considerable attention over the past few years. Much of this interest stems from the utility of the cocrystallisation process in affecting the properties (melting point, conductivity, dissolution rate etc...) of a material without changing its intrinsic chemical structure. An example of such an application is in the pharmaceutical industry, where the cocrystallisation process has been used to obtain solid forms of active pharmaceutical ingredients (APIs) with enhanced[179] physical properties. Salt preparation has always been the traditional way of modifying the physical properties of an API, especially when factors such as solubility[66] or dissolution rate demand salification or when a polymorph or solvate screen has shown that the API does not have any other available solid forms. Salts, solvates and cocrystals are all multi-component solid forms, and, despite the lack of a concensus[12-14] on what actually constitutes a cocrystal, most people would agree that both solvates and cocrystals differ from salts in that they consist of neutral molecules that are chemically distinct. The term *cocrystal* (section 1.2.1) will be used[15] for a multi-component solid form consisting of only neutral molecules, *salt* (section 1.2.1) if any pair of molecules are ionised and *disordered solid form* where the crystallography does not unambiguously place the proton at one atomic site.

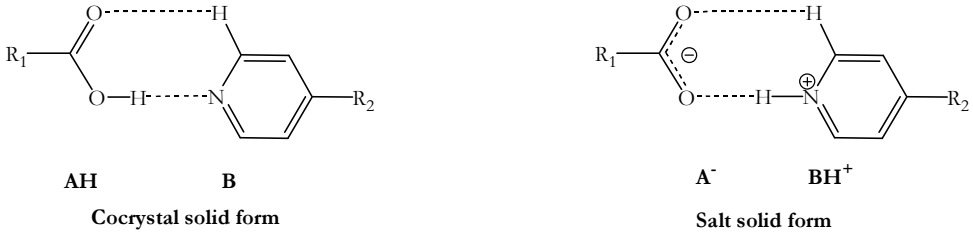
Although the literature contains many examples[14,154,160] of successful cocrystallisation experiments leading to a solid form with the expected hydrogen bonding motif as dictated by the principles of crystal engineering[180], predicting the exact three dimensional structure of the solid form resulting from such experiments is a challenging task. The solid form resulting from experiments targeting salts is perhaps even more challenging to predict because of the tendency for salt solid forms to have unpredictable lattice compositions. Aakeröy[181] has surveyed a set of 85 crystal structures consisting of both salt and cocrystal solid forms in an attempt to qualitatively understand the effects of “simple” proton transfer. All 85 multi-component solid forms were crystallised using stoichiometric amounts of a carboxylic acid and an *N*-heterocyclic base. In the study, 45 % of the salt structures were found to display a



Scheme 4.1: Molecular structures of systems used in the experimental cocrystallisation screens. **1**= Pyridine, **2**= 4-dimethylaminopyridine, **3**= Maleic acid, **4**= Fumaric acid, **5**= Phthalic acid, **6**= Isophthalic acid and **7**= Terephthalic acid.

a lattice with an unpredictable chemical (in the form of solvates) or stoichiometric composition, while just 5 % of cocrystal solid forms were found to deviate from the expected lattice composition. In this study[181], the expected stoichiometry was based on the number of mutually complementary hydrogen bond donor and acceptor sites on the acid and base. The tendency for salt solid forms to have unpredictable lattice compositions is likely to affect the endeavours of crystal engineers and theoreticians interested in the *a priori* prediction of their crystal structures.

The key physicochemical property that determines the propensity for salt formation in a solution containing an acid and base is the difference in pK_a between the conjugate acid of the base and the acid, ΔpK_a. Conclusions drawn from measured pK_a values are however valid, only under the solution equilibrium conditions at which they were determined, and these empirical parameters should strictly speaking not be used to predict or rationalise solid state behaviour. Despite this fact, they are commonly[182,183] employed for predicting solid state molecular ionisation states in crystallisation experiments involving an acid and base. When ΔpK_a is sufficiently large, salt formation is very likely, and there have been many different proposals[66,71,184] for the minimum ΔpK_a required to be confident of a salt. By contrast, in the range $0 < \Delta pK_a < 3$, experimental evidence[70,185] shows that crystallisation may result in a salt, cocrystal or disordered solid form with partial proton transfer. These difficulties in empirically predicting the molecular ionisation states for systems capable of forming either a salt or cocrystal provide a motivation for developing computational approaches as a means of



Scheme 4.2: The neutral (left), $COOH\cdots N_{arom}$, and ionic (right), $COO^{-}\cdots H\text{-}N_{arom}^{+}$, forms of the carboxylic acid-pyridine heterosynthon, both shown as part of the common $R_2^2(7)$ motif in graph set notation[30]. The same heterosynthon may be drawn using a single-point $D_1^1(2)$ graph set.

testing our understanding of the factors that determine the crystallisation outcomes.

A systematic screen for multi-component solid forms of pyridine, **1**, and 4-dimethylaminopyridine (DMAP), **2**, was performed using the set of dicarboxylic acids, **3**-**7**, shown in Scheme 4.1. The pK_a value of DMAP, **2**, implied that experiments involving **2** and each of the dicarboxylic acid coformers, **3**-**7**, would lead exclusively to salt solid forms. The pK_a of pyridine, **1,** relative to the same acids does not clearly predict the formation of a salt[185], and indeed the known solid forms[186] are a 1:1 salt (**IYUPAT**) for phthalic acid (**5**), a 2:1 cocrystal (**IYUNOF**) for terephthalic acid (**7**), and a disordered solid form (**IYUPEX**) for isophthalic acid (**6**). This set of acids and bases was chosen because of the importance[70,180] of the carboxylic acid-pyridine heterosynthon (Chapter 3, section 3.4.4), $COOH\cdots N_{arom}$, in the supramolecular synthesis of cocrystal solid forms. Indeed, a recent survey[155] of the Cambridge Structural Database[29] (CSD) has shown that in the absence of competing hydrogen bond donor and acceptor groups, the $COOH\cdots N_{arom}$ heterosynthon has a 98 % occurrence rate among a set of 126 crystal structures containing both the carboxylic acid and pyridine moieties. The ionic form of the carboxylic acid-pyridine heterosynthon, $COO^{-}\cdots H\text{-}N_{arom}^{+}$,[155] can also result from the crystallisation of simple pyridines and carboxylic acids, and the two are qualitatively distinguished (Scheme 4.2) by the position of the acidic proton between the nitrogen and oxygen atoms. Thus, the structures of the solid forms found in the screens could be used in computational modelling to assess the structural effects of the position of the acidic proton.

4.2 Method

4.2.1 Crystallisation screens

Cocrystallisation experiments involving **1** (*Acros Organics*, 99 % pure) and the dicarboxylic acids, **3-7**, were performed by dissolving each of the acid coformers in excess pyridine. The resulting solution was filtered to remove any undissolved acid and the excess pyridine was allowed to evaporate at various temperatures (-5, 5 and 25 °C). Solution crystallisation experiments involving **2** (*Alfa Aesar*, 99 % purity) and the same acids, **3-7**, were performed by briefly grinding stoichiometric amounts of acid and base using pestle and mortar and dissolving the resulting powder in the minimum amount of methanol (*Fisher Scientific*, analytical grade) needed to dissolve 0.6 g of solute. Other solvents were screened (Table 4.1) for use in the crystallisation experiments but methanol was found to be the most suitable since it dissolved all starting solid reagents (**2-7**) with the exception of terephthalic acid. Solvent evaporation was performed at the temperatures previously stated for pyridine. Automated grinding experiments were also performed as part of the screen on **2,** using a *Retsch MM200* mixer mill, equipped with 10 mL capacity stainless steel grinding jars and two 5 mm stainless steel grinding balls per jar. Each grinding experiment was performed at a frequency of 30 Hz for 30 minutes. Stoichiometric (1:1) amounts of acid and base were used in both neat and solvent drop grinding[187] experiments. In all cases the combined mass of solutes in the grinding jar did not exceed 0.4 g. Solvent drop grinding experiments were performed by adding 4 drops of methanol to the stoichiometric mixture of acid and base.

4.2.2 Single crystal X-ray diffraction

Single crystal X-ray diffraction experiments were performed on a Bruker AXS SMART APEX CCD diffractometer equipped with a Bruker AXS Kryoflex open flow cryostat [graphite monochromated Mo-K_α radiation (λ=0.71073 Å)]. Data integration and final unit cell parameters were obtained using SAINT+.[188] For all structures, absorption corrections were applied by a semi-empirical approach using SADABS[189], and the crystal structures were solved by direct methods using SHELXS-97[190]. All non-hydrogen atom positions were located using difference Fourier methods as implemented in SHELXL-97[191]. For structures **I** and **VI** (Table 4.3), all H atom positions were

located from the difference Fourier map and freely refined. For **III** (Table 4.3), all H atom positions were fixed at idealised positions and refined using a riding model. For structures **II**, **IV** and **V** (Table 4.3), the methyl protons were fixed at idealised positions and refined using a riding model, while all other H atom positions were located from the difference map and freely refined. Packing diagrams were produced using Mercury CSD 2.0[145] and the images rendered with POV-Ray[192]. Root mean square deviations ($RMSD_{15}$) for the overlay of the non-hydrogen atoms in the 15 molecule co-ordination spheres of two structures were calculated using the packing similarity feature[115] as implemented in Mercury CSD 2.0. Single molecule root mean square deviations ($RMSD_1$) were calculated using the program OptimalPaste[193].

4.2.3 Powder X-ray diffraction

Powder X-ray diffraction (PXRD) data were collected on a Stoe StadiP transmission geometry diffractometer using Ge <111> monochromated Cu $K_{\alpha 1}$ radiation (λ=1.54056 Å) operating at 40 kV and 30 mA. Diffraction patterns were collected from a sample flame sealed in a 0.5 mm diameter borosilicate glass capillary and measured with a linear position sensitive detector (nominal aperture 4.5° 2θ), which was scanned from 5 to 40° 2θ in steps of 0.2° 2θ with a count time of 140 s per step at room temperature. The scan was repeated, compared and checked for consistency, and the two scans added together to create a single summed data set with data binned in steps of 0.02° 2θ. For the mixture of **VI** and **VII** observed (Table 4.3) following neat grinding experiments, the proportion of each solid form was determined and lattice parameters refined using the Reitica[194] Rietveld refinement program. The data were not of sufficient resolution to allow the atomic coordinates to be varied. The same machine geometry was used to obtain data for the variable temperature study on the mixture of **VI** and **VII** (Table 4.3). The aim of this study was to investigate any facile thermally induced transformations between **VI** and **VII** (Table 4.3). An Oxford Instruments Cryojet was used to reach the desired temperature and data were collected in the temperature range 100 to 400 K in steps of 50 K. Data were collected for 60 minutes at each temperature with a ramping time of approximately 10 minutes between temperatures. Scans were performed from 5 to 40º 2θ in steps of 0.2º 2θ with a count time of 20 s per step.

4.2.4 Computational modelling

All well characterised salt and cocrystal solid forms shown in Table 4.3 were subjected to rigid body lattice energy minimisations using the experimentally determined molecular/ionic conformations or *ab initio* optimised molecular/ionic conformations calculated at the MP2/6-31G(d,p) level of theory using GAUSSIAN03[119]. The effects of acidic proton position were investigated by editing the experimental crystal structure in Mercury CSD 2.0[114] to give the desired sp^3 hybridised O-H or sp^2 hybridised N^+-H in the generalised hydrogen bond, N···H···O. If the observed solid form is a cocrystal, the hypothetical acidic proton position leads to a salt and vice versa. All rigid body lattice energy minimisations were performed using DMACRYS[89]. The model intermolecular potential used a distributed multipole analysis of the MP2/6-31G(d,p) *ab initio* charge density obtained using GDMA2.2[195] to model the electrostatic interactions, and an empirical *exp*-6 atom-atom potential for dispersion-repulsion effects. The force fields tested are denoted FIT(H_N), FIT(H_O, H_N) and W99[107]. Further details about the origin of these force fields and the intermolecular potential parameters used can be found in Chapter 2. In the results for the rigid body lattice energy minimisations presented later, several abbreviations are used to refer to the calculated lattice energy minima. The *Exp(Obs)MinExp* refers to the calculated lattice energy minimum using the observed molecular/ionic conformations and observed acidic proton position while the *Exp(Hyp)MinExp* is the calculated lattice energy minimum with the hypothetical acidic proton position but using the observed molecular conformations. The experimental bondlengths to hydrogen were elongated to standard[196] neutron values of 1.015 Å for O-H, 1.009 Å for N^+-H and 1.083 Å for C-H in the calculations leading to both the *Exp(Obs)MinExp* and *Exp(Hyp)MinExp* structures. The *Exp(Obs)MinOpt* and *Exp(Hyp)MinOpt* lattice energy minima were calculated using *ab initio* optimised molecular/ionic conformations at the observed or hypothetical acidic proton positions respectively.

The effects of proton transfer on the crystal energies was investigated by periodic quantum mechanical calculations using the PBE density functional and 6-31G(d,p) basis function using the CRYSTAL06[197] code. For systems (Table 4.3) **II**, **IYUPAT**, **IYUPEX** and **IYUNOF** a series of lattice energy minimisations were performed at different N···H distances ranging from 0.85-1.70 Å with all other atomic positions optimised. The space group and cell parameters were fixed at their experimental values.

All electronic structure calculations with the CRYSTAL06 code were performed by Dr. Panos Karamertzanis.

4.3 Experimental screen for salt and cocrystal solid forms of simple pyridines

4.3.1 Overview of results

Experiments involving **1** and the acid coformers **3-7** led to both salt and cocrystal solid forms. The combination of **1** and the acid coformers **5** and **7**, led to the crystal structures of the same salt and cocrystal solid forms reported by Elsegood[186]. The CSD reference codes for these structures (Table 4.3) are **IYUPAT**[186] and **IYUNOF**[186] respectively. As a consequence of the low solubility of **6** in pyridine, crystals suitable for single crystal X-ray diffraction experiments could not be grown when **6** was cocrystallised with **1.** Instead the previously published pyridine isophthalic acid cocrystal, **IYUPEX**[186] , was used in the computational modelling work. In **IYUPEX**, the acidic proton is disordered across the N···O hydrogen bond vector, with a refined site occupancy ratio of 42%:58% (N:O), in favour of the cocrystal solid form. Inspection of the geometrical details of the non-hydrogen atoms shows that the pyridine base has a C9-N1-C13 angle of 119.73 °. The C-O bondlengths of the acid are 1.23 Å (C7-O1) and 1.29 Å (C7-O2) respectively. These geometric parameters are in line with the finding of a cocrystal major product from the refinement of the crystallographic position for the disordered acidic proton. Cocrystallisation of **1** and **3** led to an in-situ base catalysed isomerisation of **3** to **4** and the resulting solid form was the same pyridine fumaric acid (2:1) cocrystal, **I**, obtained from experiments involving **4** and **1.** Cocrystallisation of DMAP, **2,** with each of the acid coformers, **3-7**, led to a salt solid form as confirmed by single crystal X-ray diffraction experiments. In all cases the 1:1 or 2:1 stoichiometry (cation:dianion) expected from the ratio of hydrogen bond donors to acceptors was found in the salt, with the exception of the experiment involving **2** and **4**. Here salts of the form 2:1:1 (**VI**) and 2:1:2 (**VII**) were concomitantly crystallised and the stoichiometric ratios correspond to the molar ratios of 4-dimethylaminopyridinium, fumarate and fumaric acid in the crystallographic asymmetric unit. The results of solubility screens for DMAP, **2**, and the acid coformers **3-7** are summarised in Table 4.1. Table 4.2 gives a

Solubilities* (mg/mL)						
Solvent	**DMAP (2)**	**Maleic acid (3)**	**Fumaric acid (4)**	**Phthalic acid (5)**	**Isophthalic acid (6)**	**Terephthalic acid (7)**
Acetone	83.33(65)	125.00(79)	-	-	-	-
Ethanol	625.00(4)	11.00(79)	-	-	-	-
Methanol	963.33(56)	388.33(79)	28.33(79)	125.00(79)	5.67(79)	-
Toluene	38.33(79)	-	-	-	-	-
Acetonitrile	57.50(79)	19.33(79)	-	-	-	-
Ethyl Acetate	21.67(79)	-	-	-	-	-
Propan-2-ol	59.00(79)	-	-	-	-	-
Water	80.83(79)	483.33(79)	-	-	-	-

***** **:** All quoted solubilities are the arithmetic mean of two experimental measurements. The quoted errors were estimated from knowledge of the inherent accuracy in the instruments used to measure the mass of solute and volume of solvent.

- **:** If the solubility of a substance was < 3 mg/mL in a given solvent, it was considered insoluble and will have no entry in the above table.

Table 4.1: Estimated solubilities of the solid reagents **2-7** in the range of solvents screened. The solvent of choice for growing single crystals in experiments between **2** and each of the acids **3-7** was methanol.

Pyridine (1)					
	SE (25 °C)	**SE (5 °C)**	**SE (-5 °C)**		
Maleic acid (3)	**I**	**I**	**I**		
Fumaric acid (4)	**I**	**I**	**I**		
Phthalic acid (5)	**IYUPAT**	**-**	**-**		
Isophthalic acid (6)	**-**	**-**	**-**		
Terephthalic acid (7)	**IYUNOF**	**IYUNOF**	**IYUNOF**		
4-dimethylaminopyridine (2)					
	SE (CH_3OH, 25 °C)	**SE (CH_3OH, 5 °C)**	**SE (CH_3OH, -5 °C)**	**NG (30 mins, frequency 30 Hz)**	**SDG (30 mins, frequency 30 Hz)**
Maleic acid (3)	**V**	**V**	**V**	**V**	**V**
Fumaric acid (4)	**VI**	**VI, VII**	**VI**	**VI, VII**	**VI, VII**
Phthalic acid (5)	**II**	**II**	**II**	**II**	**II**
Isophthalic acid (6)	**III**	**III**	**-**	**III**	**III**
Terephthalic acid (7)	**IV**	**IV**	**IV**	**IV** *	**IV** *

SE : Solvent evaporation experiment. Where the solvent is not explicitly given in parentheses, the liquid medium used to perform the crystallisation was also the base. The temperature of the crystallisation experiment is given in parentheses.

NG : Neat grinding experiment. Stoichiometric (1:1) amounts of acid and base were used. The operating conditions of the grinding mill are given in parentheses.

SDG : Solvent drop grinding experiment. Equipment and amounts of material are the same as for the NG experiment. The only difference being that 4 drops of methanol were added to each sample prior to placing the grinding jar in the mill. The operating conditions of the grinding mill are given in parentheses.

- : Crystals unsuitable for single crystal X-ray diffraction experiments.

*** :** Identification based on comparing the experimental Powder X-Ray Diffraction (PXRD) data of the sample with that simulated from the single crystal structure of 4-dimethylaminopyridinium terephthalate, **IV**. The majority of the diffraction peaks in the experimental PXRD pattern of the sample have 2θ values that match those simulated from the single crystal structure of **IV**. A few peaks in the ground sample are however not accounted for by the crystal structure of **IV**.

Table 4.2: Experimental matrix showing the solid forms obtained under the conditions used in the cocrystallisation screens on **1** and **2**. **I**= Pyridine fumaric acid (2:1) cocrystal, **II**= 4-dimethylaminopyridinium phthalate, **III**= 4-dimethylaminopyridinium isophthalate, **IV**= 4-dimethylaminopyridinium terephthalate, **V**= 4-dimethylaminopyridinium maleate, **VI**= 4-dimethylaminopyridinium fumarate-fumaric acid (2:1:1), **VII**= 4-dimethylaminopyridinium fumarate-fumaric acid (2:1:2). The structures **IYUPAT** and **IYUNOF** have been reported by Elsegood[186].

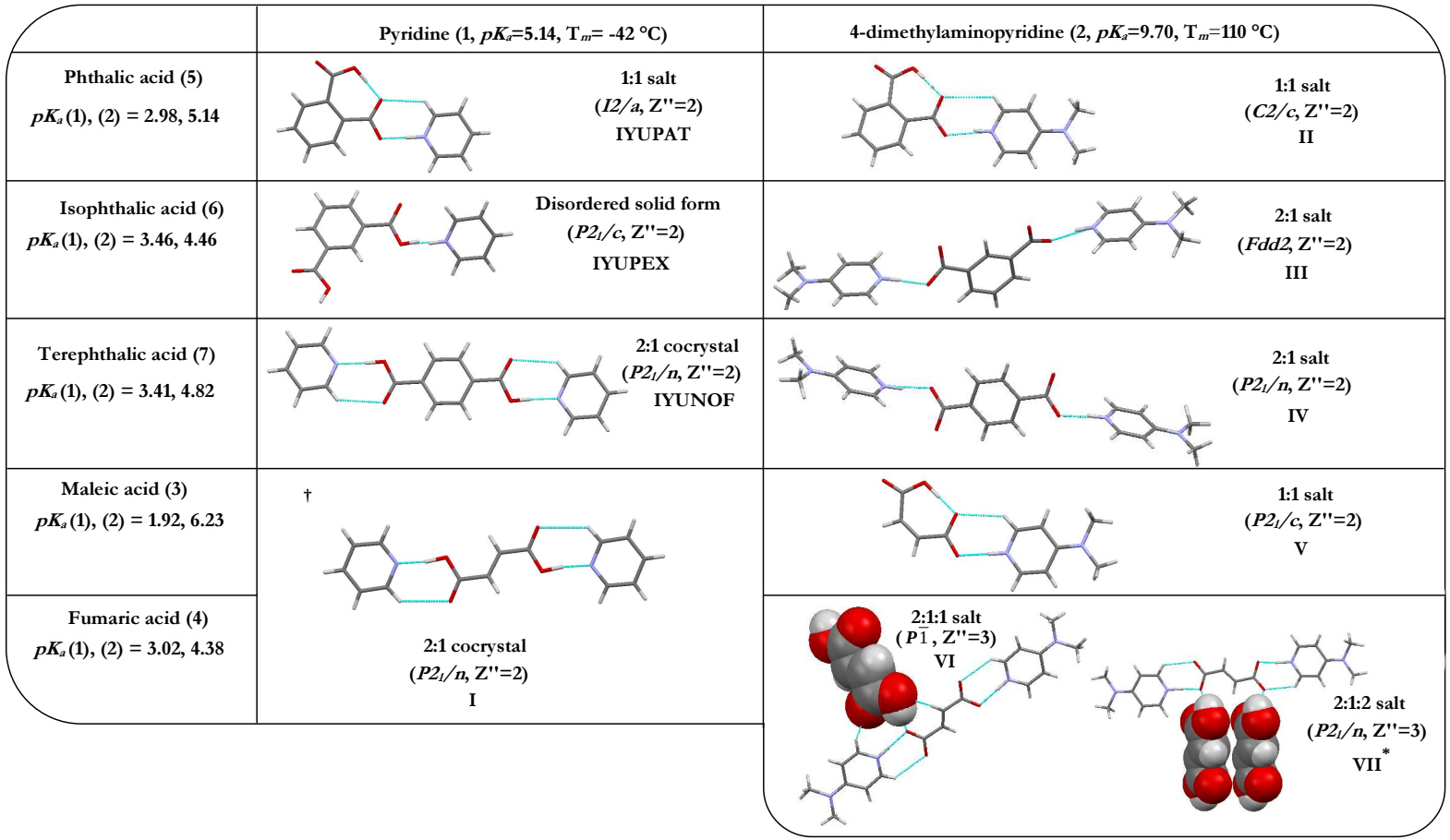


Table 4.3: Summary table depicting the results from the screens on **1** and **2**. The six character code under some of the multi-component solid forms are CSD references of previously published structures[186]. Structures **I-VII** were found as part of this work. The quoted pK_a values are literature[198,199] values that have been corrected for activity effects (with the exception of that for **2**) and all have been determined in aqueous solutions. Thermodynamic pK_a values of **1**-**7** in non-aqueous media were not found in the literature. Z′ refers to the total number of crystallographically non-equivalent molecules/ions in the asymmetric unit[200]. T_m is the melting point. † *In-situ* base catalysed isomerisation of **3** to **4** was observed when **3** was cocrystallised with **1** (see text for details). * The molecular and crystal structure of **VII** is confirmed by crystallography, but due to the low quality of the diffraction data, no significance can be attributed to the metric parameters.

summary of the observed solid forms from the screens on **1** and **2** as a function of the experimental crystallisation technique used, while Table 4.3 shows a pictorial representation of these solid forms. Crystallographic parameters for the novel solid forms found in the screens are given in Table 4.4 while a table of hydrogen bonding parameters for the solid forms studied can be found in Table 4.5.

4.3.2 Pyridine fumaric acid (2:1)

A combined differential scanning calorimetry and thermogravimetric experiment on crystals of pyridine fumaric (2:1), **I,** revealed that the cocrystal decomposes at 68 °C with the loss of two molar equivalents of pyridine. In the crystal structure of **I**, the fumaric acid molecule lies on an inversion centre with only half the molecule contained in the crystallographic asymmetric unit (Figure 4.1), along with a whole pyridine molecule. The structure displays an $R_2^2(7)$ heterodimer motif between pyridine and fumaric acid. In addition to taking part in the heterodimer motif, each oxygen atom on fumaric acid is involved in a long range intermolecular interaction with a pyridine C-H bond. In the packing diagram (Figure 4.2), this leads to localised networks consisting of two pyridine and two fumaric acid molecules that form a ring motif of graph set $R_4^4(24)$.

Given that maleic acid, **3**, has the opposite configuration around the double bond when compared to the acid in the crystal structure of **I**, it was unclear whether the crystallisation of **I** from pyridine and maleic acid was a result of an *in-situ* base catalysed isomerisation of **3** to **4** or a consequence of impurities of **4** being present in the commercial sample of **3**. Solution ^{1}H NMR experiments on **3** (Figure 4.3) using a mixture of $CDCl_3$ and DMSO-d_6 as solvent, showed that the symmetry equivalent olefinic protons display a chemical shift of 6.2 ppm in agreement with the spectrum of **3** supplied by Sigma-Aldrich. According to the reference spectrum of **4** supplied by the same company (CAS #: 110-17-8), the *trans* protons of **4** are expected at a chemical shift of 6.8 ppm. The absence of a peak at this chemical shift in the ^{1}H NMR spectrum of **3**, meant that if any impurities of **4** are present in the commercial sample of **3**, they are below the limits of detection of the instrument used in the ^{1}H NMR experiment.

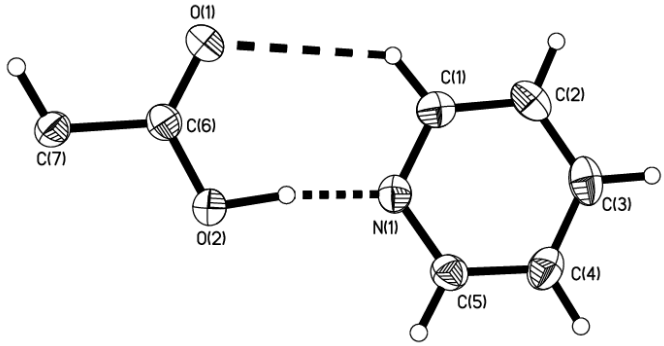


Figure 4.1: The asymmetric unit of the pyridine fumaric acid (2:1) cocrystal, **I**. Displacement ellipsoids are drawn at the 50 % probability level and hydrogen atoms are shown as spheres of arbitrary radii.

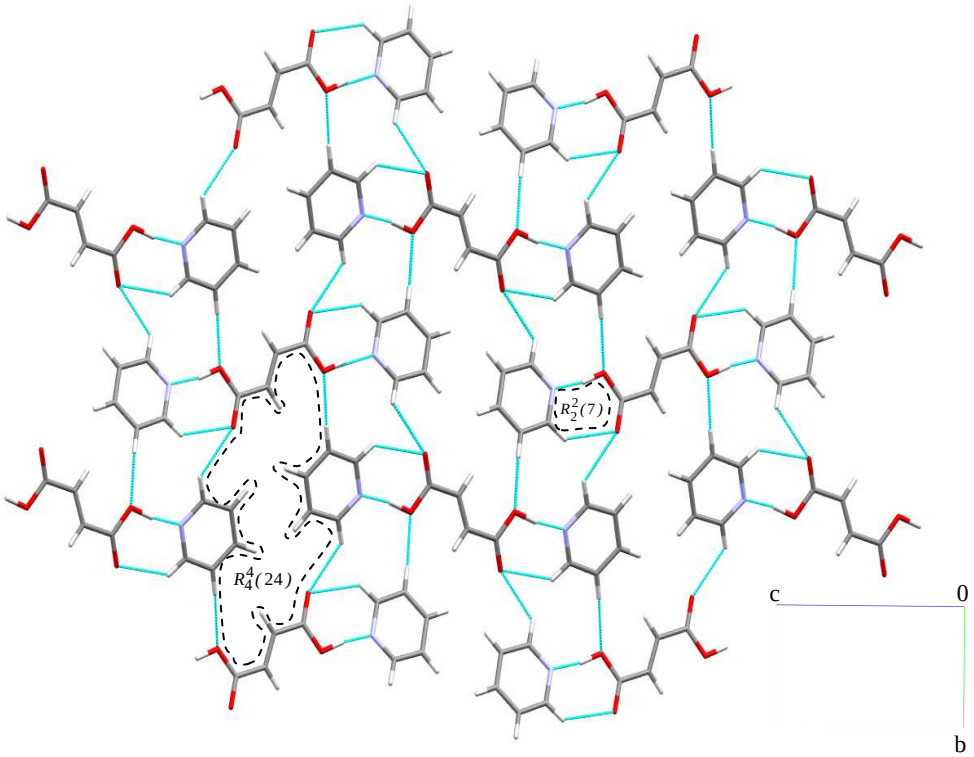


Figure 4.2: Packing diagram for the pyridine fumaric acid cocrystal, **I**. The packing diagram shows the heterodimer, $R_2^2(7)$, and ring, $R_4^4(24)$, hydrogen bonding motifs found in the crystal structure of **I**.

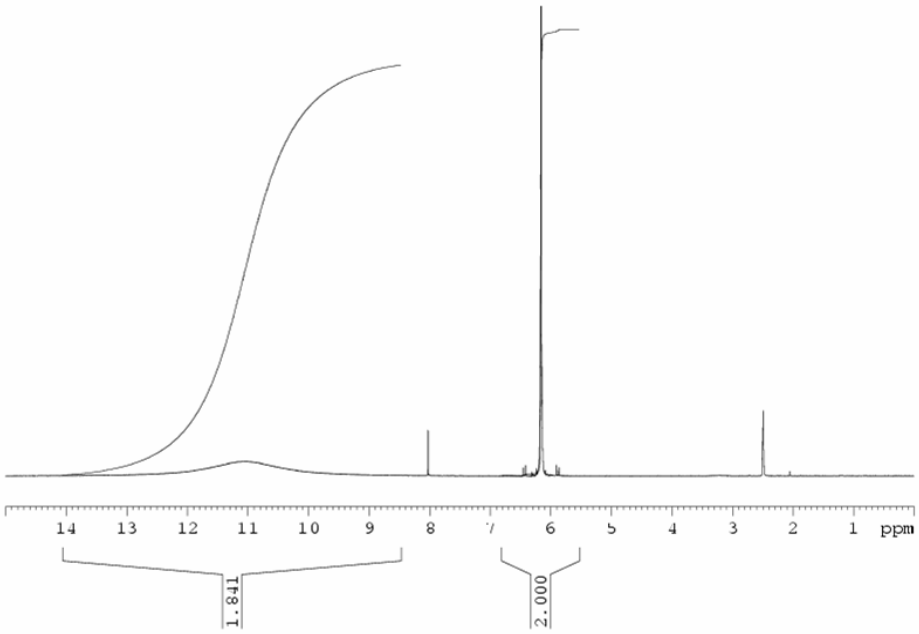


Figure 4.3: ^{1}H NMR spectrum ($CDCl_3$/DMSO-d_6) of the commercial sample of maleic acid, **3**, used in the crystallisation experiments. The *cis* protons come at a chemical shift of *ca.* 6.2 ppm, in agreement with the reference spectrum supplied by Sigma-Aldrich (CAS #: 110-16-7).

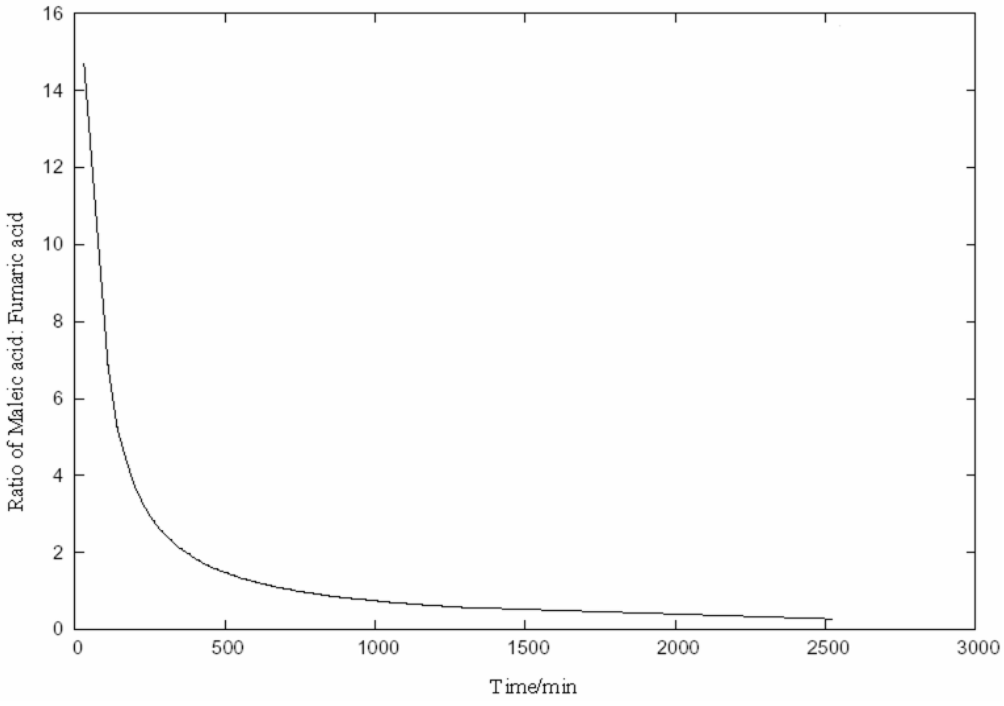


Figure 4.4: Kinetics of the base catalysed isomerisation of **3** to **4** observed when pyridine is cocrystallised with maleic acid. The plot was derived by observing changes in the ^{1}H NMR spectrum of maleic acid dissolved in deuterated pyridine (C_5D_5N).

The transformation of maleic acid, **3**, to fumaric acid, **4**, is a well known isomerisation process that proceeds catalytically in the presence of acid and at temperatures in excess of 100 °C[201]. Fumaric acid isomerises to maleic acid with a reported heat of isomerisation[202] of 5.43 kcal mol^{-1}. Chatterjee and Rao[203] have reported that cocrystallisation of **3** with 4,4′-bipyridine proceeds to give a cocrystal containing 4,4′-bipyridine and **4** in a 2:1 ratio. The authors interpreted this to be a consequence of a base (4,4′-bipyridine) catalysed isomerisation of **3** to **4**. The isomerisation was found to be solvent dependent in their studies, with only dimethylformamide (DMF) and dimethylsulfoxide (DMSO) aiding the isomerisation of **3** to **4** but other solvents such as methanol hindering it. The authors postulated that the DMF/DMSO solvent is responsible for breaking up the internal hydrogen bonding found in the solid state structure of **3**[204]. Once this is done, the 4,4′-bipyridine molecule, being a good nucleophile, was thought to add onto the double bond of the free maleic acid, thereby forming a 'zwitterionic species'. This zwitterion was then thought to isomerise to the thermodynamically more stable fumaric acid, followed by the elimination of the 4,4′-bipyridine molecule. The role of 4,4′-bipyridine in the isomerisation was confirmed by the lack of isomerisation when **3** was refluxed in DMF or DMSO in the absence of a base. An investigation of the mechanism involved in the crystallisation of **I** was not undertaken as part of this work but we did investigate the rate of isomerisation when maleic acid is dissolved in deuterated pyridine.

Deuterated pyridine (99 %, GOSS Instruments Ltd), 1 mL, was used to dissolve milligram quantities of **3** and the resulting solution was used in a ^{1}H NMR experiment conducted over a period of 44 hours. The resulting spectra were analysed and the intensity ratio of the peaks corresponding to the *cis*/*trans* olefinic protons used to infer the ratio of maleic:fumaric acid in solution as a function of time (Figure 4.4). Over a period of just 4 hours, 88 % of **3** was observed to isomerise to the thermodynamically more stable **4**. By contrast, the appearance of diffraction quality single crystals of the pyridine fumaric acid cocrystal, **I**, took one week. This rapid rate of isomerisation, in comparison to the slow processes of crystal nucleation and growth, explains why none of the crystals sampled corresponded to a pyridine maleic acid cocrystal when **3** is cocrystallised with **1**.

	I	II	III	IV	V	VI
Formula	$C_{14}H_{14}N_2O_4$	$C_{15}H_{16}N_2O_4$	$C_{11}H_{13}N_2O_2$	$C_{11}H_{13}N_2O_2$	$C_{11}H_{14}N_2O_4$	$C_{11}H_{14}N_2O_4$
Crystal system	Monoclinic	Monoclinic	Orthorhombic	Monoclinic	Monoclinic	Triclinic
Space group	$P2_1/n$	$C2/c$	$Fdd2$	$P2_1/n$	$P2_1/c$	$P\bar{1}$
a (Å)	3.8195(6)	22.130(3)	18.070(3)	6.9285(7)	13.127(3)	7.4036(13)
b (Å)	10.2834(15)	8.7473(11)	31.806(5)	15.7989(16)	7.599(2)	8.1567(14)
c (Å)	17.067(3)	15.3325(19)	6.8941(10)	9.6739(10)	12.576(3)	10.1469(17)
α (°)	90	90	90	90	90	81.381(3)
β (°)	91.436(3)	111.525(2)	90	110.646(2)	114.237(4)	89.228(3)
γ (°)	90	90	90	90	90	70.789(3)
V (Å^3)	670.15(17)	2761.0(6)	3962.3(10)	990.92(17)	1143.9(5)	571.68(17)
Z	2	8	16	4	4	2
T (K)	150(2)	150(2)	150(2)	150(2)	150(2)	150(2)
F(000)	288	1216	1744	436	504	252
D_{calc} (g cm^{-3})	1.359	1.387	1.376	1.376	1.383	1.384
μ (mm^{-1})	0.101	0.102	0.096	0.096	0.106	0.107
Crystal size* (mm^3)	$0.50 \times 0.25 \times 0.22$	$0.50 \times 0.25 \times 0.07$	$0.50 \times 0.25 \times 0.12$	$0.50 \times 0.25 \times 0.06$	$0.50 \times 0.25 \times 0.07$	$0.50 \times 0.25 \times 0.19$
Reflns collected	5581	11416	8310	8227	6290	4902
Unique reflns (R_{int})	1594 (0.0253)	3279 (0.0356)	1299 (0.0448)	2372 (0.0246)	2624 (0.0447)	2583 (0.0233)
GOOF on F^2	1.028	1.034	1.054	1.056	1.043	1.076
R_1 [$F^2 > 2\sigma(F^2)$]	0.0408	0.0505	0.0396	0.0492	0.0608	0.0463
wR_2 (all data)	0.1067	0.1287	0.1039	0.1333	0.1396	0.1213
Largest difference map features (e Å^{-3})	0.216, -0.200	0.323, -0.260	0.273, -0.230	0.367, -0.284	0.247, -0.259	0.268, -0.233

Table 4.4: Crystallographic data for the solid forms **I-VI**. * The crystals were cut to these dimensions from larger blocks.

Structure	Solid form	ΔpK_a (1)	ΔpK_a (2)	Interaction	D-H / Å	H···A / Å	D···A / Å	∠OHN / °
I	2:1 cocrystal	*2.12*	*0.76*	O2-H7···N1	1.07(2)	1.53(2)	2.5880(13)	171.3(19)
IYUPAT	1:1 salt	*2.16*	-	N(1)-H(1)···O(4) O(2)-H(2)···O(3)	1.03(2) 1.04(2)	1.53(2) 1.38(2)	2.553(2) 2.4244(18)	176.9(18) 174(2)
IYUPEX	1:1 cocrystal/1:1 salt (disordered solid form)	*1.68*	-	O(2)-H(1)···N(1) N(1)-H(1X)···O(2)	0.97(6) 0.89(7)	1.57(6) 1.65(8)	2.5402(14) 2.5402(14)	171(3) 175(5)
IYUNOF	2:1 cocrystal	*1.73*	*0.32*	O(2)-H(2)···N(1)	1.00(2)	1.63(2)	2.6286(15)	175.8(18)
II	1:1 salt	6.72	-	N(2)-H(9)···O(4) O(2)-H(16A)···O(3)	0.94(2) 0.838(10)	1.74(2) 1.589(15)	2.6679(19) 2.413(2)	168.4(19) 167(5)
III	2:1 salt	6.24	5.24	N(2)-H(9)···O(1)	0.88	1.75	2.589(2)	158.9
IV	2:1 salt	6.29	4.88	N(2)-H(9)···O(4)	0.96(2)	1.64(2)	2.5847(16)	169.5(19)
V	1:1 salt	7.78	-	N(2)-H(9)···O(1) O(3)-H(14)···O(2)	0.95(2) 1.07(3)	1.74(2) 1.36(3)	2.690(2) 2.435(2)	178(2) 179(3)
VI	2:1:1 salt	6.68	5.32	N(2)-H(9)···O(3) O(2)-H(15)···O(3)	0.99(2) 0.94(2)	1.71(2) 1.61(2)	2.6975(17) 2.5566(15)	173(2) 177.6(19)

Table 4.5: Tabulated ΔpK_a[198,199] and the N-H···O and O-H···O hydrogen bonding parameters for all solid forms studied. D and A are hydrogen bond donor and acceptor atoms. $\Delta pK_a = pK_a$ (RNH^+) - pK_a (RCO_2H), with ΔpK_a (1) and ΔpK_a (2) calculated using the pK_a for the first and second ionisation of the acid (Table 4.3) respectively. The solid forms highlighted in grey are those that fall in the range $0 < \Delta\ pK_a < 3$ where empirical evidence[185] suggests it is possible to crystallise a salt, cocrystal or solid form with properties intermediate between the two.

4.3.3 4-dimethylaminopyridinium phthalate

The 4-dimethylaminopyridinium phthalate salt, **II**, crystallises in the monoclinic space group, *C2/c*, with a 1:1 molar ratio of cation and anion in the asymmetric unit (Figure 4.5). Adopting the notation of van Eijck and Kroon[200] to describe the number of crystallographically non-equivalent molecular/ionic units (Z″) in the asymmetric unit, the phthalate salt would be described as having a structure with Z″=2. The cation and anion interact via an analogous $R_2^2(7)$ heterodimer motif as found in the crystal structure of pyridinium phthalate, **IYUPAT**[186]. The phthalate anion displays an O2-H16···O3 intramolecular hydrogen bond (Figure 4.5) of graph set $S(7)$. The proton involved in this intramolecular hydrogen bond was found to be disordered across the O2···O3 hydrogen bond vector with a refined site occupancy ratio of 47(6)%:53(6)% for the two positions. The O-H bondlength was fixed to the standard X-ray distance of 0.84(1) Å. The $S(7)$ intramolecular hydrogen bond is a common feature in both ionised[205] and non-ionised[206] 1,2-dicarboxylic acids and is also found in the crystal structure of pyridinium phthalate, **IYUPAT**.

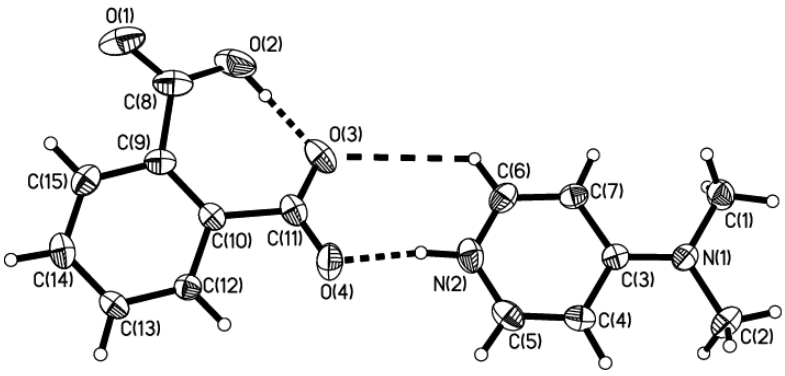


Figure 4.5: Asymmetric unit of 4-dimethylaminopyridinium phthalate, **II**. Displacement ellipsoids are drawn at the 50 % probability level and hydrogen atoms are shown as spheres of arbitrary radii. Only one component of the intramolecular O2···H16···O3 poton disorder is shown.

4.3.4 4-dimethylaminopyridinium isophthalate

4-Dimethylaminopyridinium isophthalate, **III**, crystallises in the orthorhombic space group, *Fdd2*, with half the isophthalate dianion as well as a complete 4-dimethylaminopyridinium cation contained in the crystallographic asymmetric unit. In **III**, the dianions are arranged in infinite columns (Figure 4.6) parallel to the *c*-axis.

4.3.5 4-dimethylaminopyridinium terephthalate

4-Dimethylaminopyridinium terephthalate, **IV**, crystallises in the space group $P2_1/n$. The asymmetric unit contains half the terephthalate dianion as well as a complete 4-dimethylaminopyridinium cation, making it a 2:1 salt. Hydrogen bonded tapes consisting of cation (C) and anion (A) molecules follow the sequence ACCACCA and propagate parallel to the *b*-axis (Figure 4.7). The tapes interact to form ring motifs of graph set $R_4^3(12)$ and $R_4^4(46)$. The terephthalate dianions are arranged in columns similar to that found in **III**.

4.3.6 4-dimethylaminopyridinium maleate (1:1)

Solution crystallisation experiments involving a 1:1 stoichiometric ratio of **2** and **3** led to crystals of 4-dimethylaminopyridinium maleate, **V**. Both neat and solvent drop grinding experiments also led to this solid form, **V**. Single crystal X-ray diffraction experiments showed that **V** crystallises in the monoclinic space group, $P2_1/c$, with a 1:1 molar ratio of cation and anion in the asymmetric unit. The maleate anion interacts with the 4-dimethylaminopyridinum cation via an $R_2^2(7)$ heterodimer motif. The anion also displays the $S(7)$ intramolecular hydrogen bond found in the phthalate anion of **II**. The heterodimer interactions propagate into two tapes. Using mean planes defined by representative $R_2^2(7)$ heterodimer interactions, the tapes were found to lie at an angle of 59.13 º degrees with respect to one another, and are linked by C-H···O contacts.

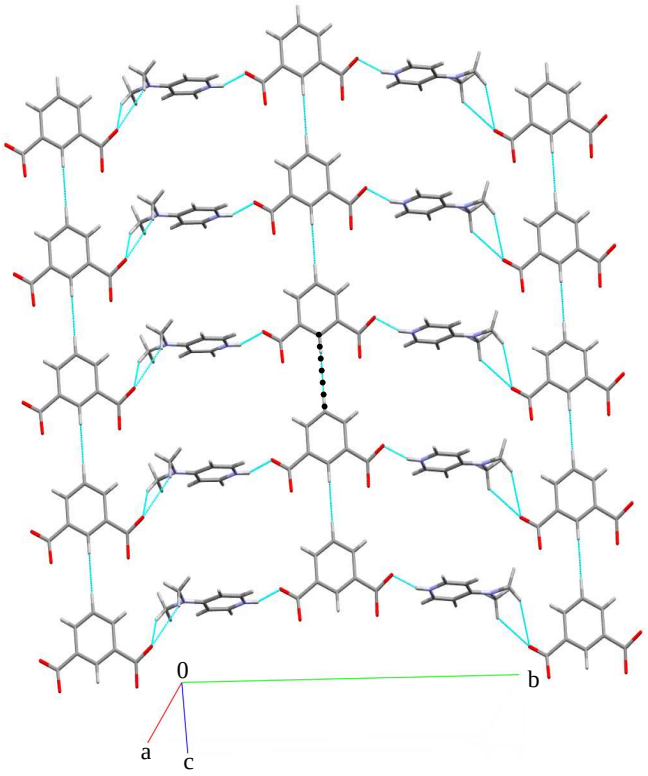


Figure 4.6: Illustration of the columns of isophthalate dianions found in 4-dimethylaminopyridinium isophthalate, **III**. The shortest separation (illustrated by dotted line) between adjacent isophthalate ions corresponds to a C12···C8 distance of 4.136 Å. Very similar columns of terephthalate ions are seen in **IV**, with the corresponding C11···C11 distance of 4.162 Å.

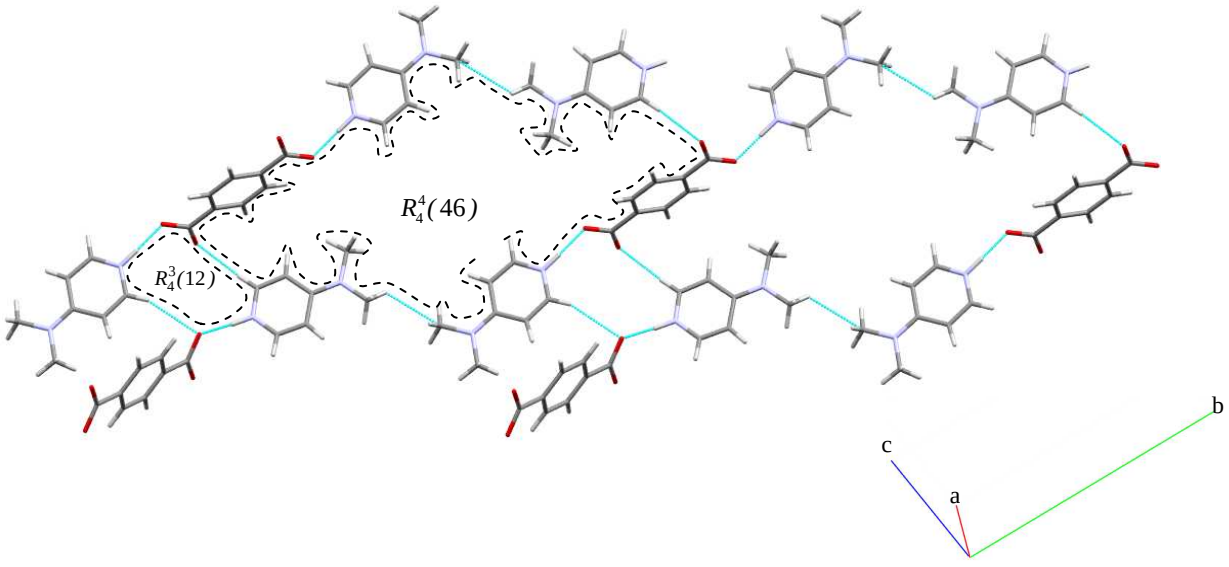


Figure 4.7: Hydrogen bonded tapes found in the structure of 4-dimethylaminopyridinium terephthalate, **IV**. The $R_4^3(12)$ and $R_4^4(46)$ ring motifs are shown by the dotted lines.

4.3.7 4-dimethylaminopyridinium fumarate-fumaric acid

Solution crystallisation experiments involving a 1:1 stoichiometric ratio of fumaric acid, **4**, and 4-dimethylaminopyridine, **2**, led exclusively to crystals of block morphology when the solvent evaporation was performed at a temperature of 25 °C. Repeating the experiment at 5 °C, led to a mixture of block and plate like crystals (Table 4.2), with the latter accounting for a minor fraction of the total amount of crystalline material. The crystals of block morphology were characterised as 4-dimethylaminopyridinium fumarate-fumaric acid (2:1:1), **VI**, while single crystal X-ray diffraction experiments performed on the plate like crystals revealed them to be 4-dimethylaminopyridinium fumarate-fumaric acid (2:1:2), **VII**. Mechanical grinding experiments using a 1:1 stoichiometric ratio of **2** and **4** also led to a sample containing a mixture of **VI** and **VII**. The mixture was obtained following neat grinding and solvent drop grinding experiments (Figure 4.8). Rietveld refinement (Figure 4.9) of the powder X-ray diffraction data of the mixture produced by neat grinding showed that **VI** accounted for 93(1) % of the total amount of material.

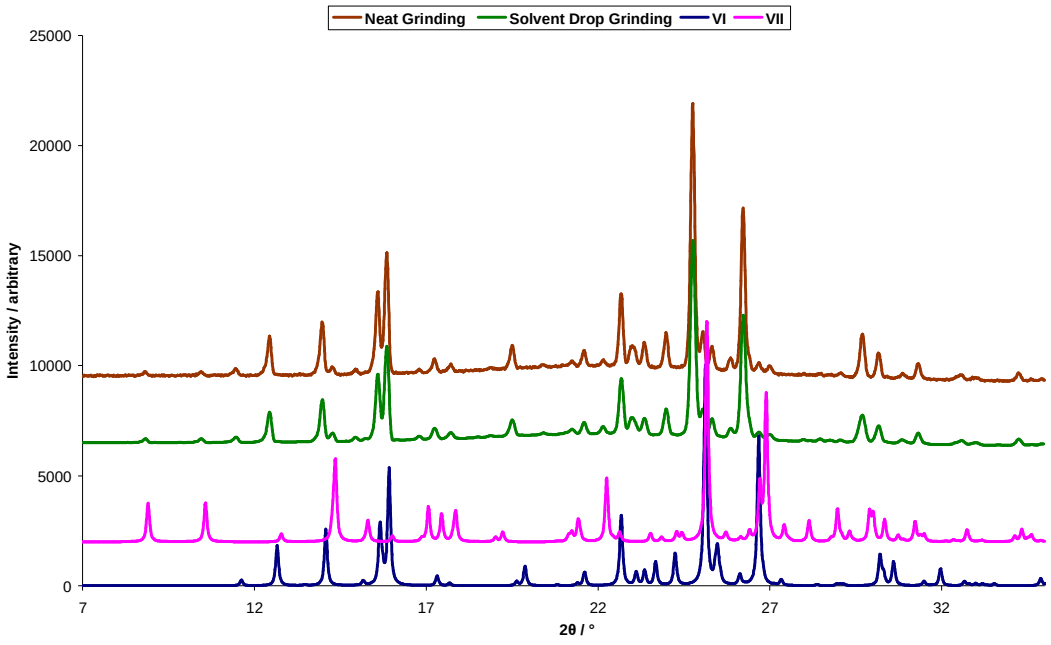


Figure 4.8: An overlay of the simulated powder patterns from the crystal structures of **VI** (blue) and **VII** (pink) with the experimental powder pattern of a stoichiometric (1:1) mixture of **2** and **4** that had been subjected to either neat (brown) or solvent drop grinding (green) conditions. The non-uniform differences in the peak positions of the simulated and experimental powder patterns reflect the observed anisotropy in the thermal expansion behaviour of the unit cells for **VI** and **VII**.

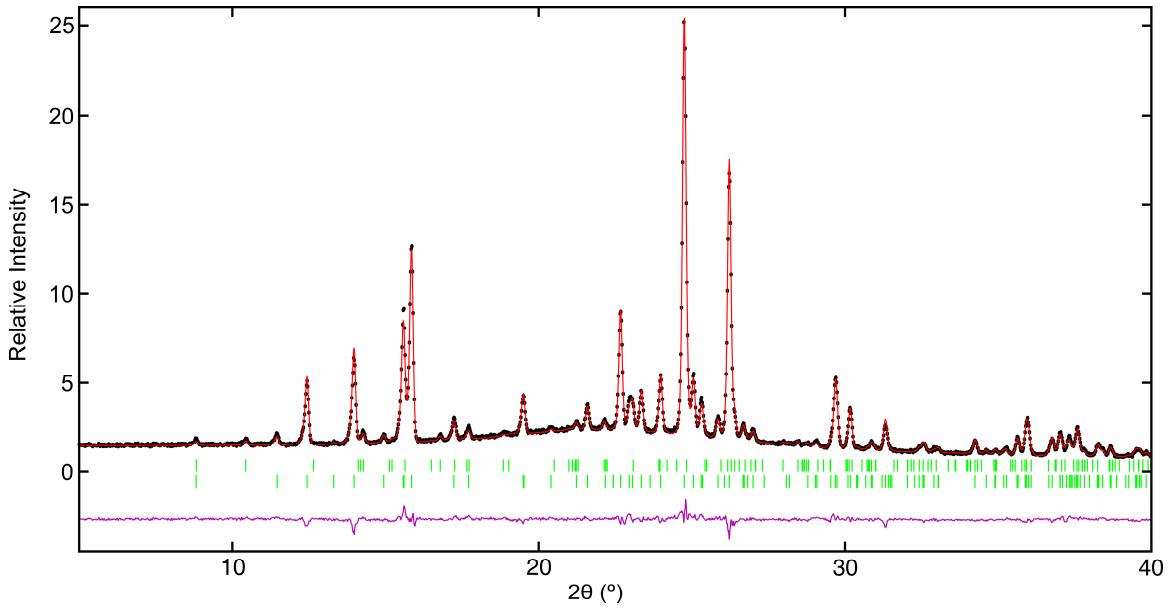


Figure 4.9: Rietveld plot showing the fit (R_{wp}=3.77 %) between the transmission PXRD data of the mixture of **VI** and **VII** produced by neat grinding with a model consisting of the cell parameters derived from the single crystal structures. Black dots indicate raw data while the red line indicates the calculated model. Upper tick marks are the 2θ positions for the hkl reflections of **VI** while the lower tick marks represent those of the minor component, **VII**. The difference pattern is shown in purple.

4-dimethylaminopyridinium fumarate-fumaric acid (2:1:1), **VI**, crystallises in the triclinic space group, $P\bar{1}$, with both the fumaric acid and fumarate molecules lying on inversion centres. By contrast, the 2:1:2 salt, **VII**, has only the fumarate molecule lying on an inversion centre and crystallises in the monoclinic space group, $P2_1/n$. The poor quality diffraction data collected for **VII** was judged to be sufficient in confirming the molecular and crystal structure, but was not used in the rigid body lattice energy minimisations presented later because of the sensitivity of the results to the accuracy of the input experimental structure. The main difficulty with the refinement of structure **VII** was an unusually high isotropic displacement parameter for the N^+-H acidic proton and a freely refined N^+-H bondlength of 1.23(4) Å. The ratio of observed to unique reflections in the refined structure was also found to be 42 %. Despite these difficulties, comparison of the symmetry unique C-O distances of the fumarate dianion as well as the C-N-C angle of the 4-dimethylaminopyridinium cation, with the values typical of COOH···N_{arom} and COO^-···H-N_{arom}^+ motifs[28], confidently identified **VII** as a salt rather than a cocrystal. Variable temperature PXRD data collected on a mixture of **VI** and **VII** showed no facile thermally induced transformations between the two solid forms in the temperature range 100-400 K. The variable temperature PXRD experiments took a total

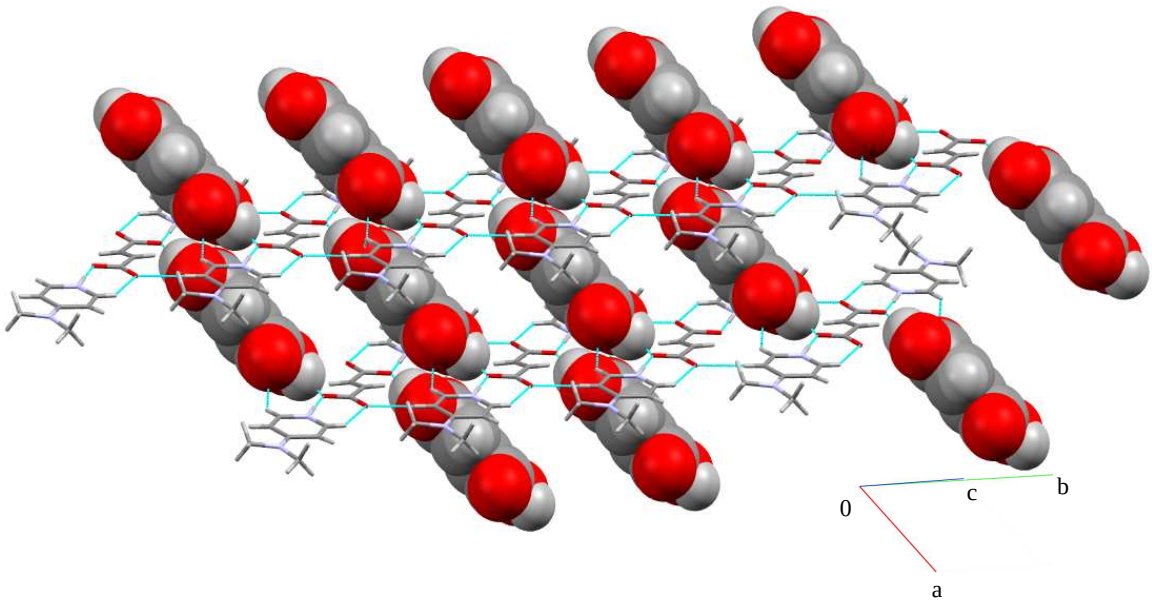


Figure 4.10: Illustration of the sheet structure found in **VI**. The fumaric acid molecules (shown using space filling models) pack between sheets of hydrogen bonded 4-dimethylaminopyridinium and fumarate ions.

of 8 hours to perform. For solid forms **VI** and **VII**, the Rietveld refinement of the lattice parameters as a function of temperature showed anisotropy in the thermal expansion behaviour of the unit cells. For **VI**, the anisotropy of the lattice expansion is illustrated by the following changes in lattice parameters observed between the temperatures of 150 K and 300 K: $(\Delta a / a)\times 100 = 1.07\,\%$, $(\Delta b / b)\times 100 = 0.78\,\%$ and $(\Delta c / c)\times 100 = -0.09\,\%$. Within the same temperature range, the anisotropy of the lattice expansion in **VII** is illustrated by the following changes in the lattice parameters: $(\Delta a / a)\times 100 = 1.07\,\%$, $(\Delta b / b)\times 100 = 1.36\,\%$ and $(\Delta c / c)\times 100 = 0.22\,\%$. For both crystal structures, the anisotropy is illustrated by the negligible changes in the *c*-lattice parameter when compared to the strong dependence of the *a*- and *b*-lattice parameters with temperature. It is this anisotropy in the thermal expansion behaviour of the unit cells in **VI** and **VII** that cause the non-uniform differences in the peak positions (greater at high 2θ) of the simulated PXRD patterns from structure determinations at 150 K and the observed PXRD pattern for the mixture at room temperature (shown in Figure 4.8). Such non-uniform difference in powder patterns due to anisotropy in the thermal expansion behaviour of unit cells has previously been noted[207].

In the asymmetric unit of **VI**, a complete 4-dimethylaminopyridinium cation interacts with the symmetry unique portion of the fumarate dianion via an $R_2^2(7)$ heterodimer motif (Scheme 4.2). In the extended crystal structure, the inversion operator leads to a trimer consisting of two cations and one dianion. These trimers are arranged in sheets (Figure 4.10), through a C-O···HC interaction of one fumarate oxygen with a 4-dimethylaminopyridinium cation on an adjacent trimer. The other crystallographically unique fumarate oxygen forms a hydrogen bond to a fumaric acid molecule, which links the sheets. By contrast, **VII** contains the same $R_2^2(7)$ heterodimer motif between cation and anion, but both fumarate oxygen atoms are hydrogen bonded to fumaric acid molecules, thereby preventing the formation of the sheet structure found in **VI**.

The inclusion of fumaric acid in the crystal structures of fumarate salts is not uncommon[208]. In our experiments, the large difference in the relative solubilities of fumaric acid and 4-dimethylaminopyridine (Table 4.1, methanol solvent) suggested that using an equimolar amount of acid and base to target the salt would not guarantee a solid form free of fumaric acid. That is why the crystallisation experiment was repeated using equal concentrations of acid and base but this led exclusively to crystals of **VI**. The CSD currently contains 10 salts with fumaric acid *syn-anti*[208,209] hydrogen bonded to fumarate anions as in **VI** and **VII**. Previous work by Ballabh[210] *et al*, which looked at the hydrogen bonding motifs in four ammonium carboxylate salts including benzylammonium fumarate, also found the *syn-anti* arrangement to be preferred for fumarate salts that contain neutral fumaric acid. Haynes and Pietersen[208] have commented that the use of the fumarate *syn* lone pair of electrons for hydrogen bonding with the cation, as directed by the favourable $R_2^2(7)$ dimer motif, means that the fumarate anion is left only with the *anti* lone pair of electrons for interaction with fumaric acid. An explanation as to why neutral fumaric acid should crystallise with fumarate salts in the first place has not been forthcoming from the work of either Haynes[208] or Ballabh[210].

4.3.8 Conclusions on experimental screens

The crystallisation screens on pyridine, **1**, and 4-dimethylaminopyridine, **2**, were performed using the same set of dicarboxylic acid coformers, **3-7**. The combination of

these dicarboxylic acid coformers with **2** led to salt solid forms (**II-VII**). Suitable single crystals of solid forms **II-VII** were grown using methanol solvent. The solid forms that result from the cocrystallisation of **1** with the acid coformers **5-7** have been previously reported by Elsegood[186]. Experiments involving **1** and **5** led to a known salt (**IYUPAT**[186]) and those involving **1** and **7** led to a known cocrystal (**IYUNOF**[186]). In both cases, there were no signs of polymorphism according to the results of the various solution and solid state grinding experiments performed. Suitable single crystals could not be grown when **1** was cocrystallised with **6**, and instead the reported pyridine isophthalic acid (**IYUPEX**[186]) structure was used in the computational modelling work. The presence of a disordered carboxylic acid proton in **IYUPEX** meant that while the cocrystal appeared as the major, 58(6) %, product from the crystallographic refinement, there was also the possibility of refining the crystal structure as a salt, 42(6) %. Both **3** and **4** were cocrystallised with pyridine in separate crystallisation experiments. Both experiments led to the crystal structure of the same pyridine fumaric acid cocrystal, **I**. Solution ^{1}H NMR experiments later confirmed that **3** undergoes an *in situ* base (pyridine) catalysed isomerisation to **4**. The rate of isomerisation was found to be much faster than the combined rate for the processes of crystal nucleation and growth. This explains why none of the crystals sampled corresponded to a pyridine maleic acid cocrystal, when pyridine was cocrystallised with maleic acid.

4.4 Modelling the structural and energetic effects of acidic proton position

4.4.1 Choice of suitable force field

The suitability of the FIT(H_N), FIT(H_O, H_N) and W99 force fields (Chapter 2) for modelling the salt and cocrystal solid forms shown in Table 4.3 was investigated. The indicators used to judge the suitability of the force field were the deviation in the intermolecular contact distance between oxygen and nitrogen in the generalised hydrogen bond O···H···N, changes in the unit cell parameters following lattice energy minimisation as well as the root mean square deviation ($RMSD_{15}$) for overlaying the experimental and lattice energy minimised structures. Details regarding the last two indicators as well as the % changes in the cell parameters are summarised in Table 4.6

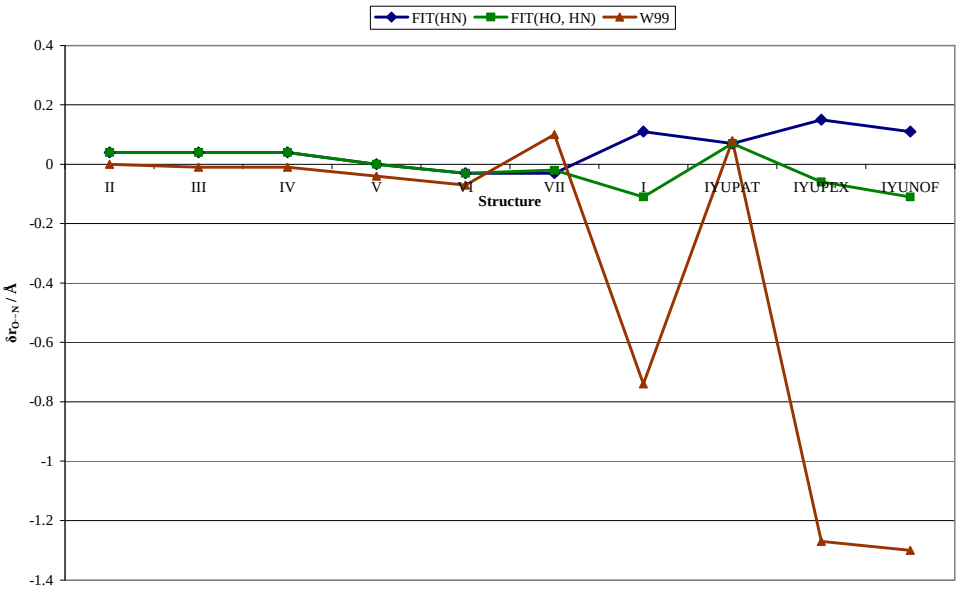


Figure 4.11: Plot of the change in the oxygen-nitrogen intermolecular distance, $\delta r_{O\cdots N}$, upon lattice energy minimising the experimental structure with the observed molecular/ionic conformations. The error is defined according to $\delta r_{O\cdots N} = r_{O\cdots N}[Exp(Obs)MinExp] - r_{O\cdots N}[Experimental]$.

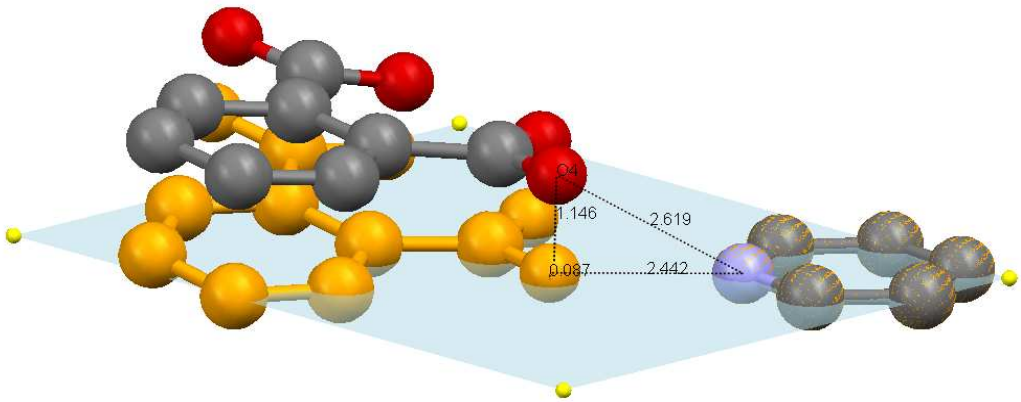


Figure 4.12: The overlay of the *Exp(Obs)MinExp* of **IYUPAT** (coloured by element) with the *Exp(Hyp)MinExp* (coloured orange) structure. The overlay is with respect to the pyridine moieties in the two structures. In the *Exp(Obs)MinExp* structure, the phthalate and pyridinium ions are not coplanar, giving a deviation of 1.146 Å for O4 (of the phthalate carboxylate group) from the mean plane of the pyridinium cation when compared with the observed value of 0.156 Å. This slippage effect is not observed in the *Exp(Hyp)MinExp* structure, where the acid and base are roughly coplanar as illustrated by the mean plane deviation of 0.087 Å for O4. All hydrogen atoms have been removed for clarity. The lattice energy minima were calculated using the FIT(H_O, H_N) force field.

Name of Structure	Intermolecular Potential	Intermolecular heteroatom distance (Å)		a (Å) [% Δa]	b (Å) [% Δb]	c(Å) [% Δc]	α (°) [% Δα]	β (°) [% Δβ]	γ (°) [% Δγ]	Volume (Å^3) [% ΔV]	U (kJ mol^{-1})	F	RMSD$_{15}$ (Å)
		O····N	O···O										
Pyridinium phthalate [**IYUPAT**]	-	2.55	-	22.0544(19)	3.7779(3)	26.481(2)	90	94.598(2)	90	2199.3(3)	-		-
Exp(Obs)MinExp	DMA + FIT(H_N)	2.62	-	6.45	0.28	-4.40	0	-0.62	0	2.13	-571.53	180.89	0.399
	DMA + FIT(H_O, H_N)	2.62	-	6.49	0.23	-4.43	0	-0.55	0	2.08	-571.80	182.85	0.400
	DMA + W99	2.63	-	6.55	1.57	-5.42	0	-0.95	0	0.03	-558.31	238.89	0.416
Exp(Obs)MinOpt	DMA + FIT(H_N)	2.68	-	6.75	1.90	-5.41	0	0.19	0	2.87	-556.61	238.43	0.400
	DMA + FIT(H_O, H_N)	2.68	-	6.78	1.82	-5.42	0	0.22	0	2.80	-556.88	239.92	0.414
	DMA + W99	2.69	-	6.74	3.13	-6.21	0	0.26	0	3.20	-545.38	298.89	0.424
Pyridine isophthalic acid [**IYUPEX**]	-	2.54	2.67	3.7969(3)	17.2355(15)	17.2272(15)	90	90.374(2)	90	1127.35(16)	-		-
Exp(Obs)MinExp	DMA + FIT(H_N)	2.69	2.92	-1.67	3.52	2.70	0	-1.25	0	4.53	-192.23	47.87	0.204
	DMA + FIT(H_O, H_N)	2.48	2.81	-0.85	2.21	1.45	0	-1.43	0	2.79	-211.68	21.61	0.179
	DMA + W99	1.27	6.26	*	*	*	*	*	*	*	*	*	*
Exp(Obs)MinOpt	DMA + FIT(H_N)	2.69	2.95	-1.61	4.04	2.80	0	-2.13	0	5.20	-177.65	62.43	0.244
	DMA + FIT(H_O, H_N)	2.51	2.85	-0.91	2.82	1.90	0	-3.24	0	3.72	-192.16	45.40	0.239
	DMA + W99	0.59	3.05	*	*	*	*	*	*	*	*	*	*
Pyridine terephthalic acid [**IYUNOF**]	-	2.63	-	9.9855(7)	7.3212(5)	11.9778(9)	90	113.144(2)	90	805.17(10)	-		-
Exp(Obs)MinExp	DMA + FIT(H_N)	2.74	-	2.52	-4.08	3.74	0	0.55	0	1.54	-126.98	48.67	0.235
	DMA + FIT(H_O, H_N)	2.52	-	0.21	-3.44	2.66	0	-2.08	0	1.00	-140.29	40.12	0.271
	DMA + W99	1.33	-	*	*	*	*	*	*	*	*	*	*
Exp(Obs)MinOpt	DMA + FIT(H_N)	2.75	-	2.43	-3.90	5.39	0	0.98	0	2.87	-117.19	68.83	0.298
	DMA + FIT(H_O, H_N)	2.56	-	0.36	-3.41	5.34	0	-0.45	0	2.50	-127.29	59.96	0.317
	DMA + W99	1.17	-	*	*	*	*	*	*	*	*	*	*
Pyridine fumaric acid [**I**]	-	2.59	-	3.8195(6)	10.2834(15)	17.067(3)	90	91.436(3)	90	670.15(17)	-		-
Exp(Obs)MinExp	DMA + FIT(H_N)	2.70	-	0.68	4.31	-0.56	0	3.63	0	4.11	-119.28	46.90	0.202
	DMA + FIT(H_O, H_N)	2.48	-	0.13	4.61	-2.40	0	1.78	0	2.11	-134.78	51.13	0.200
	DMA + W99	1.85	-	*	*	*	*	*	*	*	*	*	*
Exp(Obs)MinOpt	DMA + FIT(H_N)	2.69	-	0.25	6.26	0.07	0	3.52	0	6.28	-110.24	74.88	0.259
	DMA + FIT(H_O, H_N)	2.49	-	-3.10	6.46	0.79	0	-4.26	0	3.91	-122.98	195.44	0.358
	DMA + W99	0.71	-	*	*	*	*	*	*	*	*	*	*

Table 4.6: Results for the lattice energy minimisation of the experimental crystal structures obtained from the cocrystallisation screen on pyridine, **1**, as a function of the force field used. All quoted lattice energies are per mole of the pyridinium ion (**IYUPAT**) or pyridine molecule (**I**, **IYUPEX**, **IYUNOF**). The calculations highlighted in sky blue were performed using the W99 force field and led to invalid lattice energy minima with intermolecular separation between non-bonded atoms approaching values typical of covalently bonded systems.

Name of Structure	Intermolecular Potential	Intermolecular heteroatom distance (Å)		a (Å) [% Δa]	b (Å) [% Δb]	c(Å) [% Δc]	α (°) [% Δα]	β (°) [% Δβ]	γ (°) [% Δγ]	Volume (Å^3) [% ΔV]	U (kJ mol^{-1})	F	RMSD$_{15}$ (Á)
		O····N	O···O										
4-dimethylaminopyridinium phthalate [II]	-	2.67	-	22.130(3)	8.7473(11)	15.3325(19)	90	111.525(2)	90	2761.0(6)	-		-
Exp(Obs)MinExp	DMA + FIT(H_N)	2.71	-	-2.68	4.44	2.41	0	1.20	0	3.10	-551.11	57.27	0.290
	DMA + FIT(H_O, H_N)	2.71	-	-2.67	4.41	2.40	0	1.20	0	3.08	-551.19	56.76	0.289
	DMA + W99	2.67	-	-3.60	3.41	2.79	0	0.91	0	1.74	-544.74	54.31	0.301
Exp(Obs)MinOpt	DMA + FIT(H_N)	2.74	-	-0.98	2.37	4.46	0	0.83	0	5.19	-537.06	47.96	0.273
	DMA + FIT(H_O, H_N)	2.74	-	-0.99	2.32	4.48	0	0.82	0	5.16	-537.16	48.10	0.274
	DMA + W99	2.70	-	-1.67	1.51	4.21	0	0.62	0	3.51	-531.52	41.25	0.259
4-dimethylaminopyridinium isophthalate [III]	-	2.59	-	18.070(3)	31.806(5)	6.8941(10)	90	90	90	3962.3(10)	-		-
Exp(Obs)MinExp	DMA + FIT(H_N)	2.63	-	0.66	-0.61	1.75	0	0	0	1.79	-712.70	6.80	0.101
	DMA + W99	2.58	-	-0.49	-0.79	0.76	0	0	0	-0.53	-712.21	4.27	0.057
Exp(Obs)MinOpt	DMA + FIT(H_N)	2.62	-	0.80	1.54	2.53	0	0	0	4.95	-689.37	20.10	0.182
	DMA + W99	2.58	-	0.34	1.09	1.28	0	0	0	2.74	-688.52	11.37	0.212
4-dimethylaminopyridinium terephthalate [IV]	-	2.59	-	6.9285(7)	15.7989(16)	9.6739(10)	90	110.646(2)	90	990.92(17)	-		-
Exp(Obs)MinExp	DMA + FIT(H_N)	2.63	-	2.03	0	0.06	0	-1.02	0	2.83	-351.89	9.22	0.169
	DMA + W99	2.58	-	1.41	-0.19	-1.55	0	-1.35	0	0.60	-351.01	9.74	0.161
Exp(Obs)MinOpt	DMA + FIT(H_N)	2.63	-	2.40	1.57	-0.10	0	-0.20	0	4.06	-343.47	54.17	0.361
	DMA + W99	2.58	-	1.87	0.88	-0.75	0	0.08	0	1.93	-342.48	49.60	0.348
4-dimethylaminopyridinium maleate [V]	-	2.69	-	13.127(3)	7.599(2)	12.576(3)	90	114.237(4)	90	1143.9(5)	-		-
Exp(Obs)MinExp	DMA + FIT(H_N)	2.69	-	-0.65	1.62	0.73	0	-0.68	0	2.30	-559.53	7.44	0.112
	DMA + FIT(H_O, H_N)	2.69	-	-0.64	1.57	0.74	0	-0.63	0	2.24	-559.73	7.16	0.110
	DMA + W99	2.65	-	-1.22	1.11	0.03	0	-0.54	0	0.38	-557.48	4.87	0.089
Exp(Obs)MinOpt	DMA + FIT(H_N)	2.72	-	-0.75	2.53	1.18	0	-0.74	0	3.64	-546.77	15.40	0.166
	DMA + FIT(H_O, H_N)	2.72	-	-0.72	2.47	1.20	0	-0.68	0	3.57	-546.99	14.92	0.164
	DMA + W99	2.68	-	-1.37	2.13	0.44	0	-0.63	0	1.74	-544.25	10.55	0.142
4-dimethylaminopyridinium fumarate-fumaric acid (2:1:1) [VI]	-	2.69	2.56	7.4036(13)	8.1567(14)	10.1469(17)	81.381(3)	89.228(3)	70.789(3)	571.68(17)	-		-
Exp(Obs)MinExp	DMA + FIT(H_N)	2.66	2.69	0.93	0.47	0.51	-1.03	-0.58	-2.63	0.48	-396.76	16.84	0.138
	DMA + FIT(H_O, H_N)	2.66	2.51	-0.93	-0.32	1.00	-1.47	0.04	0.48	-0.42	-402.77	9.62	0.111
	DMA + W99	2.62	2.23	-3.64	-1.70	1.63	-2.29	1.36	6.01	-2.14	-409.83	56.84	0.274
Exp(Obs)MinOpt	DMA + FIT(H_N)	2.67	2.72	0.55	4.26	-1.22	-3.16	-5.64	0.20	2.83	-382.80	79.58	0.324
	DMA + FIT(H_O, H_N)	2.68	2.58	-0.80	3.70	-0.73	-3.20	-4.87	2.08	2.21	-387.48	64.32	0.308
	DMA + W99	2.63	2.36	-2.86	1.76	1.24	-2.65	-0.91	4.91	1.40	-390.25	51.22	0.293

Table 4.7: Results for the calculated lattice energy minima of the experimental crystal structures obtained from the screen on 4-dimethylaminopyridine, **2**. For structures **II**, **V** and **VI**, the lattice energy minima were calculated using the FIT(H_N), FIT(H_O, H_N) or W99 force fields. The FIT(H_O, H_N) force field was not used for structures **III** and **IV** because the structures do not contain any O-H protons. All quoted lattice energies are per mole of the 4-dimethylaminopyridinium cation.

and Table 4.7. Figure 4.11 shows the change in the intermolecular oxygen-nitrogen distance in hydrogen bonds of type $COOH\cdots N_{arom}$ or $COO^{-}\cdots H\text{-}N_{arom}^{+}$ upon relaxing the experimental structure. The figure shows that by contrast to the FIT(H_N) and FIT(H_O, H_N) force fields, the W99 force field underestimates (*i.e.* smaller than experimental value) the O···N intermolecular distance for all structures with the exception of structures **II** and **IYUPAT**. Minimising the crystal lattice energy of structure **II** with the W99 force field leads to no change in the intermolecular O···N distance, while both the FIT(H_N) and FIT(H_O, H_N) force fields, lead to a minor increase in the intermolecular separation of 0.04 Å. For **IYUPAT**, all the force fields seem to overestimate the O···N distance, and they all lead to a lattice energy minimum with a significant difference in the co-planarity of the pyridinium and phthalate ions when compared to the relative position of these species found in **IYUPAT**. This slippage effect is best illustrated (Figure 4.12) by considering the deviation of the phthalate O4 atom from the mean molecular plane of the pyridinium cation. In the experimental (**IYUPAT**) structure, this deviation is only 0.156 Å. By contrast, lattice energy minimisation using the FIT(H_O, H_N) force field leads to a more pronounced deviation of 1.146 Å. The most striking differences in behaviour between the FIT and W99 force fields appears to be for structures **I**, **IYUPEX** and **IYUNOF**, where the W99 force field seriously underestimates the O···N distances to the extent that the lattice energy minima calculated for these structures are invalid. This indicates a problem with the parameterisation of the W99 force field when used in conjunction with a distributed multipole electrostatic model to describe dispersion-repulsion effects in strongly hydrogen bonded systems. Day has also observed unreasonably short oxygen-oxygen intermolecular distances when using the carboxylic acid H(3) potential - as implemented in the W99[147,211] force field - to cocrystal XV (aminopyridimidine: *o*-toluic acid) in conjunction with distributed multipoles in the fourth international blind test of crystal structure prediction[212]. He found that by replacing the potential parameters (cf. Table 2.1) of H(3) for those of H(2), which represent polar protons in OH groups, more accurate hydrogen bonding geometries and distances are observed. In lattice energy minimisations involving **I, IYUPEX** and **IYUNOF**, the problem is not limited to H(3), and inspection of the structures corresponding to the calculated lattice energy minima has shown that H(1), H(4) and C(3) all seem to be performing badly when used in conjunction with a distributed multipole electrostatic model. Of the two

remaining force fields that seem suitable for modelling the salt and cocrystal solid forms of **1** and **2**, the results in Table 4.6 and Table 4.7 show that on average, the FIT(H_O, H_N) leads to lattice energy minimised structures with hydrogen bond interaction distances that are closer to the experimental values when compared to the results obtained using FIT(H_N). As a consequence, the FIT(H_O, H_N) force field was used to investigate the structural effects of acidic proton position in the rigid body modelling work.

4.4.2 Investigating the effects of ionisation state on molecular conformation

Ab initio geometry optimisations were performed for the individual molecules/ions found in the crystal structures shown in Table 4.3 (with the exception of **VII**) as part of the calculations leading to the *Exp(Obs)MinOpt* structures. The resulting optimised conformation for each molecule/ion represents the lowest energy planar conformation in the gas phase assuming no crystal packing effects. Single molecule overlays ($RMSD_1$) of the experimental (Exp) and optimised (Opt) conformations therefore reveal useful information regarding the effects of crystal packing forces on the adopted molecular/ionic conformations in each crystal structure and these have been tabulated (*Exp(Obs) vs. Opt(Obs)*, Table 4.8). In addition to this, the effect of ionisation state on the calculated *ab initio* gas phase conformational minimum was investigated by overlaying the geometry of the hypothetical ionisation state for each molecule/ion with the *ab initio* gas phase minimum (*Exp(Hyp) vs. Opt(Hyp)*, Table 4.8) of this ionisation state. For each molecule/ion, the hypothetical ionisation state is obtained by adding/removing the necessary number of protons required to change the system from an ionic to a neutral solid form or vice versa, as dictated by the molar stoichiometric ratio of species found in the crystallographic asymmetric unit. The final column in Table 4.8, *Opt(Obs) vs. Opt(Hyp)*, compares the conformations of the gas phase minima of the observed and hypothetical ionisation states.

Table 4.8 shows that the experimental conformation of the pyridine or pyridinium ions deviates very little from the lowest energy *ab initio* gas phase minimum in the crystal structures **IYUPAT**, **IYUPEX**, **IYUNOF** and **I**. This is indicated by an average $RMSD_1$ of less than 0.1 Å for the overlays in the column entitled *Exp(Obs) vs. Opt(Obs)* under the sub-heading "all atoms". By contrast, the salts of **2** (**II-VI**) all display higher $RMSD_1$ values for the 4-dimethylaminopyridinium cation when the

$RMSD_1$ / Å						
		Exp(Obs) *vs.* Opt(Obs)		Exp(Hyp)[$] *vs.* Opt(Hyp)		Opt(Obs) *vs.* Opt(Hyp)
Molecule/Ion	**Structure**	**All atoms**	**Non H atoms**	**All atoms**	**Non H atoms**	**Non H atoms**
4-dimethylaminopyridinium	**II**	0.178	0.066	0.208	0.110	0.106
	III	0.369	0.057	0.439	0.128	-
	IV	0.361	0.039	0.304	0.092	-
	V	0.105	0.026	0.167	0.091	-
	VI	0.215	0.057	0.272	0.111	-
Pyridinium	**IYUPAT**	0.101	0.019	0.096	0.029	0.041
Pyridine	**I**	0.095	0.018	0.101	0.033	-
	IYUPEX	0.097	0.022	0.106	0.024	-
	IYUNOF	0.089	0.016	0.097	0.036	-
Maleate	**V**	0.070	0.034	0.113	0.079	0.076
Fumarate	**VI**	0.144	0.065	0.155	0.087	0.093
Fumaric acid	**I**	0.136	0.056	0.149	0.092	-
	VI	0.224	0.082	*	*	-
Phthalate	**II**	0.080	0.034	0.223	0.186	0.192
	IYUPAT	0.114	0.080	0.456	0.421	-
Isophthalate	**III**	0.157	0.165	0.160	0.161	0.056
Isophthalic acid	**IYUPEX**	0.106	0.088	0.116	0.093	-
Terephthalate	**IV**	0.164	0.113	0.161	0.113	0.061
Terephthalic acid	**IYUNOF**	0.112	0.073	0.118	0.086	-

Table 4.8: Illustration of the differences between the experimental, *Exp*, and *ab initio* optimised, *Opt*, molecular/ionic conformations as a function of whether the acidic proton position is at the experimentally observed, *Obs*, or hypothetical, *Hyp*, position in the generalised hydrogen bond O···H··N. The final column compares two *ab initio* optimised conformations, one at the observed acidic proton position, *Opt(Obs)*, and the other at the hypothetical acidic proton position, *Opt(Hyp)*. [$] The *Exp(Obs)* and *Exp(Hyp)* conformations are the same since proton transfer was assumed not to be associated with any conformational changes. [*] This is a neutral molecule within a salt structure that is not involved in COO···H···N hydrogen bonds, and as a consequence does not have a hypothetical acidic proton position.

values under the same column are inspected. This is a consequence of the methyls of the NMe_2 group, which under the influence of the packing forces in the experimental crystal structures are able to rotate significantly from the position found in the *ab initio* gas phase minimum. The effect of methyl rotation in the calculated overlays of 4-dimethylaminopyridinium is apparent when comparing the $RMSD_1$ data calculated using only non-H atoms and those calculated taking into account all atom positions. The $RMSD_1$ data calculated using non-H atom positions are on average 80 % lower

than the values calculated taking into account all atom positions. For the acidic/anionic counterparts modelled, anions of 1,2-dicarboxylic acids (maleate and phthalate) tended to give *ab initio* optimised molecular/ionic conformations which were not significantly different to the experimentally observed conformations. The degree of similarity between the *Exp(Obs)* and *Opt(Obs)* conformations is largely due to the effect of the intramolecular hydrogen bond between the carboxylic and carboxylate groups and the effect this has in inhibiting significant changes in torsion angle involving these groups. With regards to the effects of ionisation state on the calculated *ab initio* gas phase conformational minima, Table 4.8 shows that there is quantitatively very little conformational difference between the pyridine and pyridinium ions. This is reflected in the similarity of the calculated $RMSD_1$ data for the *Exp(Obs) vs. Opt(Obs)* and *Exp(Hyp) vs. Opt(Hyp)* overlays ("all atoms") for the pyridine molecules in the structures **I**, **IYUPEX** and **IYUNOF**. By contrast, *ab initio* geometry optimisation of the 4-dimethylaminopyridinium cation leads to a gas phase conformational minimum that is distinctly different to that calculated for the 4-dimethylaminopyridine molecule, **2**. In the *ab initio* geometry optimisation of **2**, there is noticeable pyramidalisation of the NMe_2 group. This phenomenon is not observed when the *ab initio* optimised conformation of the cation is calculated, and this is reflected in the much lower $RMSD_1$ data for the *Exp(Obs) vs. Opt(Obs)* overlay (sub-heading "Non H atoms") for the cations when compared to data for the same type of overlay involving the hypothetical molecules of **2** shown in the column entitled *Exp(Hyp) vs. Opt(Hyp)*. The phthalate anion also seems to be sensitive to ionisation state, as judged by a relatively large $RMSD_1$ value of 0.192 Å for overlaying the *ab initio* gas phase conformational minima of the phthalate and phthalic acid (**5**) molecules.

4.4.3 Structural and energetic effects of acidic proton position in salts and cocrystals

Rigid body lattice energy minimisations were performed on the experimental solid forms using two types of acidic proton position: The experimentally observed position as well as a hypothetical position generated following removal of the acidic proton from its observed position and placing at the alternative position in the generalised hydrogen bond O···H···N, using the correct sp^2 or sp^3 hybridised N^+-H or O-H geometry. All

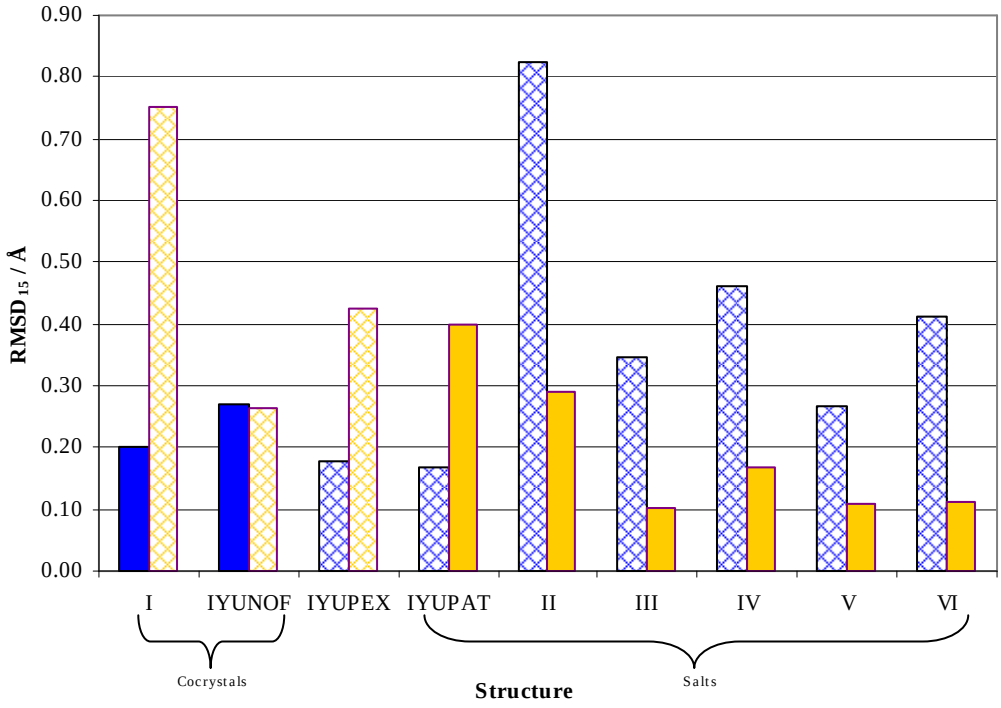


Figure 4.13: The figure shows the root mean square deviation ($RMSD_{15}$) associated with the overlay of the experimental solid form (blue for cocrystals, orange for salts) with the lattice energy minimum calculated assuming the observed (solid) covalent bond to the acidic proton, *Exp(Obs)MinExp*, or a hypothetical (hatched) covalent bond, *Exp(Hyp)MinExp*, generated by editing the structure in Mercury CSD 2.0. The acidic proton of **IYUPEX** is disordered across the N···O hydrogen bond vector and the bar graphs corresponding to both X-H (X=O or N^+) positions are shown hatched so as to indicate they are both plausible positions for the acidic proton. All lattice energy minima were calculated using the FIT(H_O, H_N) force field.

cocrystal solid forms were therefore modelled as salts and all salts modelled as cocrystals. Electronic structure calculations were performed to see the energy barrier associated with proton transfer for selected salt and cocrystal systems. Unlike the rigid body lattice energy minimisations, the electronic structure calculations were able to take into account the subtle changes in the geometrical parameters of the molecules/ions as a function of acidic proton position - most notably the C-O bondlengths of carboxylic acids and the C-N-C angle of pyridines - along the N···O hydrogen bond vector.

For the rigid body lattice energy minimisations, the *Exp(Obs)MinExp* and *Exp(Hyp)MinExp* lattice energy minima were overlaid with the experimental crystal structure for each solid form and the root mean square deviation for the overlay was computed assuming a co-ordination sphere of 15 molecules ($RMSD_{15}$). The numerical results for these overlays are given in Figure 4.13. The figure shows that for the majority of systems studied, when the salt and cocrystal positions for the acidic proton

are relaxed, the resulting lattice energy minima are structurally very different. Furthermore, with the exception of **IYUPAT** and **IYUNOF**, the calculated lattice energy minimum is more structurally related to the experimental structure when the correct acidic proton position is used than is the case when the incorrect proton position is assumed. This is judged by the lower $RMSD_{15}$ values for the solid blue and orange bars (correct proton position) in Figure 4.13 when compared to the hatched bars (incorrect proton position) of the same colours. In **IYUNOF**, both the observed and hypothetical proton positions lead to lattice energy minima which are structurally similar. In **IYUPAT**, the calculations seem to suggest that relaxing the experimental crystal structure with the hypothetical acidic proton position leads to a lattice energy minimum that is more structurally related to **IYUPAT** than is the case when the experimental structure is relaxed with the observed acidic proton position. This is a consequence of a slippage effect between the cation and anion encountered during the calculation of the *Exp(Obs)MinExp* structure (Figure 4.12). The fact that this slippage effect is not encountered in the calculation leading to the *Exp(Hyp)MinExp* structure explains why the $RMSD_{15}$ for overlaying the experimental solid form with the *Exp(Hyp)MinExp* structure is more than a factor of 2 smaller than that for overlaying the experimental solid form with the *Exp(Obs)MinExp* structure. On the whole, the results in Figure 4.13 show that if the proton is incorrectly positioned, the resulting forces are sufficient to cause quite dramatic reorientations of the rigid body molecular entities.

Figure 4.14 contrasts the relaxed N···H scans for **IYUPEX**, **IYUPAT**, **IYUNOF** and **II** from the fixed cell electronic structure calculations. With the exception of **IYUNOF**, the minimum lattice energy position for the acidic proton corresponds to a salt, although in all cases this is outside the 3σ range of the mean neutron[196] diffraction N^+-H distance for salts. For **IYUPEX**, the N^+-H distance of 1.150 Å in the minimum lattice energy structure differs markedly from the standard[196] neutron bondlength of 1.009(19) Å. Moreover, the energy penalty to elongate the N-H bond in the range 1.05-1.50 Å does not exceed 5 kJ mol^{-1}, which is consistent with the observed disorder in the acidic proton of **IYUPEX**. However, the finding that the salt solid form of **IYUPEX** is calculated to be more stable than the cocrystal is contrary to the fact that the cocrystal is observed as the major product from the crystallographic refinement of the acidic proton position. There are two points to note here. First of all, the experimental occupancy ratio of 58(6)%:42(6)% for the disordered acidic proton position being attached to the

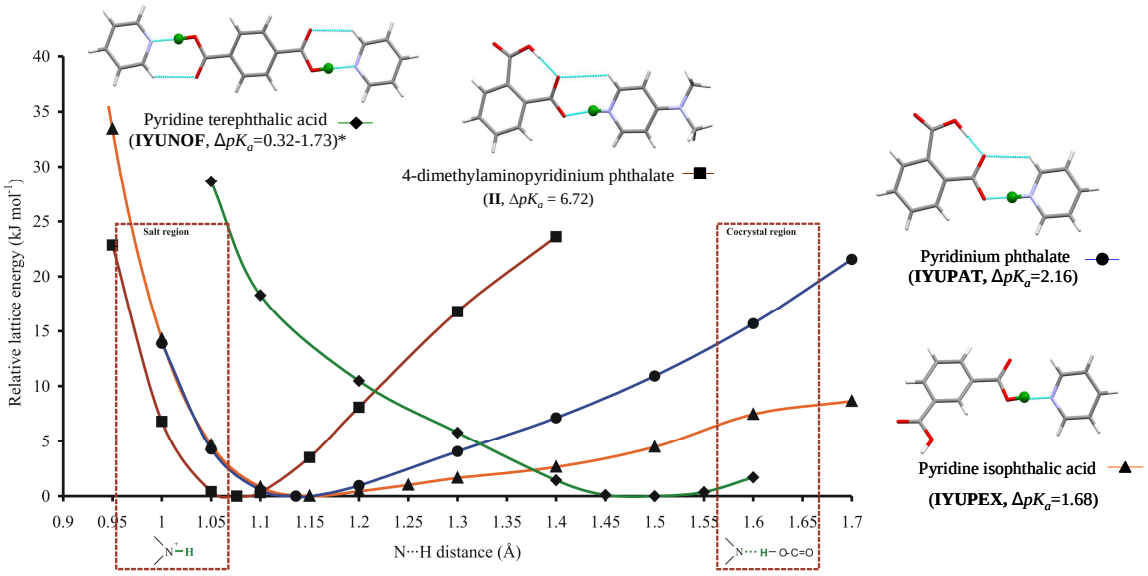


Figure 4.14: Variation in the relative lattice energy with the N···H distance of the acidic (green) proton in fixed unit cell periodic electronic structure calculations. The blue areas give the N···H distances typical of salts and cocrystals, defined by the average N_{arom}-H and O-H (in CO_2H) neutron diffraction[196] distances (±3σ). The cocrystal N···H distance is derived assuming a linear hydrogen bond with the N···O distance of 2.6286 Å of **IYUNOF**. * Maintaining the centre of symmetry in **IYUNOF** means that both N···H distances are kept identical during the electronic structure calculations, which does not correspond to the sequential ionisation found in solution. A range of 0.32-1.73 is given for ΔpK_a so as to indicate that as far as the computational calculations are concerned, the effective ΔpK_a lies between $\Delta pK_a(1)$ and $\Delta pK_a(2)$. All data points were provided by Dr. Panos Karamertzanis.

acid (cocrystal) or base (salt) is within the margin of error and in itself is insufficiently conclusive to be confident of the cocrystal over the salt. Secondly, it is important to bear in mind that the energy estimates from the modelling are true at 0 K while the crystal structure of **IYUPEX** was determined at 150(2) K. Thus it is conceivable that in light of the small energy penalty of 7.43 kJ mol^{-1} for proton transfer at 0 K, the stability ordering will reverse at 150(2) K and the cocrystal will be more stable than the salt. In Figure 4.14, the energy penalty to elongate the N-H bondlength is consistent with the expected trend on the basis of pK_a differences (*i.e.* systems with larger ΔpK_a values pay a greater penalty). The salt, 4-dimethylaminopyridinium phthalate, **II**, has a well defined N-H bondlength of 1.08 Å in the computed lowest energy structure, which is in good agreement with the experimental value[186] of 0.94 (2) Å (Table 4.5), given the underestimation of X-H bondlengths by X-ray diffraction. The potential energy well is much shallower for pyridinium phthalate (**IYUPAT**), in agreement with the smaller ΔpK_a for this acid-base pair. In **IYUNOF**, the minimum energy position for the acidic

proton corresponds to a cocrystal. The carboxylic O-H bondlength in the computed minimum lattice energy structure is 1.08 Å, while the observed distance is 1.00(2) Å. The pyridine fumaric acid system (**I)** readily optimises to the cocrystal if the fixed cell quantum mechanical lattice energy minimisation is performed from either the observed or hypothetical proton position. The minimum lattice energy acidic proton position in **VII** corresponds to a salt with the computed C-O distances (1.283 Å and 1.276 Å) of the dianion as well as the C-N-C angle (120.64 °) of the cation in agreement with the values found from the structure determination for this system and those typical of other pyridinium carboxylate salts[28].

4.4.4 Conclusions on modelling the effects of acidic proton position

The accurate location of hydrogen atom positions is essential if meaningful results are to be obtained from the computational modelling of salt and cocrystal solid forms. However this is often limited by the quality of the experimental data, and even with the best structure model from the X-ray crystallography, the systematic underestimation of X-H bondlengths is something that has to be corrected for prior to using the structure in a calculation. Sometimes, an uncertainty in acidic proton position is an indication of an underlying chemical property of the system and should not be interpreted solely as a result of poor experimental data. For example, when the relative acidity of an acid-base combination is within the range[185] $0 < \Delta pK_a < 3$, disorder of the acidic proton across the N···O hydrogen bond vector may indicate a solid form that is genuinely intermediate in property between a salt and cocrystal. Two contrasting modelling techniques were used to look at the effects of acidic proton position in the salt and cocrystal systems considered. The rigid body lattice energy minimisations performed using the correct and hypothetical acidic proton positions were intended to probe the structural effects of acidic proton position, and revealed on the whole that the resulting lattice energy minimum structure is sensitive to the location of the acidic proton. While such calculations provide an insight into the structural effects of acidic proton position, their utility in resolving suspected ambiguities in acidic proton position is limited by the accuracy of the model potential parameters used. For this reason, electronic structure methods were performed on systems **II**, **IYUNOF**, **IYUPAT** and **IYUPEX**. The

calculations correctly predict the observed solid form (salt or cocrystal) for all systems with the exception of **IYUPEX** which was found to have a range of energetically accessible N^+-H geometries, in agreement with the experimentally observed disorder. For all systems, the computed lowest energy N^+-H position lies outside the 3σ range of the mean neutron diffraction N^+-H distance for salts and that expected for cocrystals assuming a linear N···H···O hydrogen bond. The qualitative agreement of the N^+-H position in lattice energy minima for most systems with the experimental acidic proton positions suggests that such electronic structure calculations using known unit cells may be useful in providing a strong indication as to the probable acidic proton positions when there is experimental uncertainty.

4.5 Conclusions

A screen for multi-component solid forms of pyridine and 4-dimethylaminopyridine using maleic, fumaric, phthalic, isophthalic and terephthalic acids has led to five fully characterised salt solid forms (**II-VI**) of 4-dimethylaminopyridine as well as the crystal structure of a novel pyridine fumaric acid cocrystal (**I**). The previously reported[186] crystal structures **IYUPAT** and **IYUNOF** were also found in the screens, with **IYUPEX** proving difficult to grow suitable single crystals for. Maleic acid was found to isomerise too quickly in pyridine and repeated crystallisation experiments led to the same pyridine fumaric acid cocrystal, **I**. The combined solid forms from both screens illustrate the problems in empirically predicting the stoichiometry and covalent bonding of the acidic proton within the structures of salts and cocrystals. The majority of the solid forms in the limited 2×5 grid of structures were based on either the neutral or ionic form of the carboxylic acid-pyridine $R_2^2(7)$ heterosynthon (Chapter 3). If the acidic protons had not been located by crystallography, visual inspection would not have confidently assigned them as a salt or cocrystal. Nevertheless, computational modelling methods have shown that the location of the acidic proton is important for modelling the crystal structures of salts and cocrystals. This work has contributed to our understanding of the differences between organic salts and cocrystals through the results of computational modelling, which have shown that experiments to relocate the acidic proton of the salt or cocrystal lead to significant structural changes in the majority of

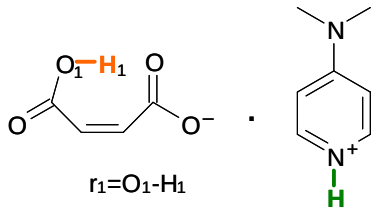
cases. The calculations have also shown that the majority of structures can be better modelled when the crystallographic designation of salt or cocrystal is used and this has important implications for *ab initio* crystal structure prediction of salts and cocrystals.

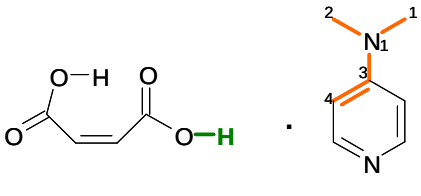
5 Crystal structure prediction of pyridinium carboxylate salts and pyridine carboxylic acid cocrystals

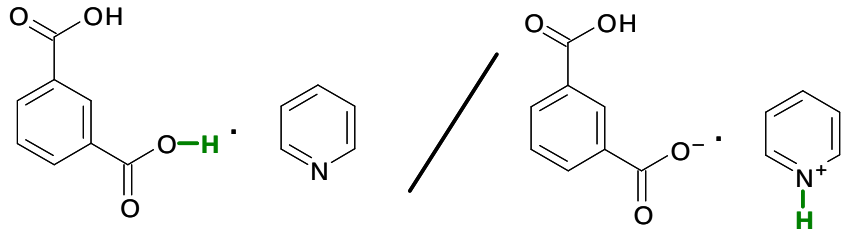
5.1 Introduction

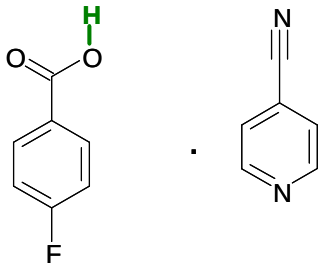
From the work presented in the last chapter, we have seen that the results of static lattice energy minimisations of the structures of pyridinium carboxylate salts and pyridine carboxylic acid cocrystals are sensitive to the assumed position of the acidic proton within the crystal. The main implication of this finding from the perspective of our modelling work is that the results of *ab initio* crystal structure prediction of salts and cocrystals will also be sensitive to the assumed covalent bonding of the acidic proton. The aims of this chapter are two fold. Firstly, we hope to validate the extent to which existing computational modelling methods can successfully predict the crystal structures of simple pyridinium carboxylate salts and pyridine carboxylic acid cocrystals. We will also investigate the sensitivity of the predicted crystal energy landscape to the position of the acidic proton in simulations that assume the molecular structures of the salt and cocrystal. The study will be limited to three types of system (Table 5.1): a salt, **V**, formed from complete proton transfer from acid to base, a disordered salt-cocrystal system, **IYUPEX**[186], and a cocrystal where the acidic proton is fully localised on the carboxylic acid, **VIII**. Structures **V** and **IYUPEX** are 1:1 multi-component systems taken from the work of Chapter 4. The lack of a suitable 1:1 cocrystal from the results of Chapter 4 meant that it was necessary to screen and subsequently characterise the crystal structure of **VIII** which is a novel cocrystal of 4-cyanopyridine and 4-fluorobenzoic acid. Under aqueous conditions, the combination of 4-cyanopyridine (pK_a=1.86)[213] and 4-fluorobenzoic acid (pK_a=4.15)[214] leads to a ΔpK_a of -2.29. This indicated the suitability of this acid-base pair for targeting a cocrystal of the two compounds since it has previously[70] been suggested that when $\Delta pK_a < 0$, crystallisation is likely to result in a cocrystal as opposed to a salt. From Chapter 4, the results of manual experimental screens have shown that there are no polymorphic forms of **V** and **IYUPEX** under the range of crystallisation conditions employed.

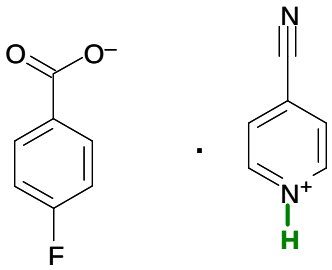
Prior to this work, theoretical studies have shown that *ab initio* crystal structure prediction of cocrystals is not only feasible[79,122,212,215], but increasingly useful[124,125] in

<table>
<tr><th colspan="2">Systems tested in computational crystal structure prediction</th></tr>
<tr><th>Experimental molecular structures</th><th>Hypothetical molecular structures</th></tr>
<tr><td> 

4-dimethylaminopyridinium maleate (V)</td><td> 

4-dimethylaminopyridine maleic acid</td></tr>
<tr><td colspan="2"> 

Pyridine isophthalic acid:Pyridinium isophthalate (58%:42%)
(IYUPEX)</td></tr>
<tr><td> 

4-cyanopyridine 4-fluorobenzoic acid
(VIII)</td><td> 

4-cyanopyridinium 4-fluorobenzoate</td></tr>
</table>

Table 5.1: Illustration of the experimental molecular structures of the pure salt, **V**, pure cocrystal, **VIII**, and disordered salt-cocrystal, **IYUPEX**, systems tested as well as the associated hypothetical molecular structures expected following transfer of the acidic proton in green. For **IYUPEX**, the proton disorder leads to the molecular structures of the salt and cocrystal according to the percentage ratios quoted. The bondlengths and torsion angles (r_1, θ_1 and θ_2) that were constrained to experimental neutron/X-ray diffraction values during *ab initio* geometry optimisation are indicated in orange.

understanding the organic solid state when complemented with the results of experimental screens. By contrast, *ab initio* crystal structure prediction of molecular carboxylate salts has remained a largely unexplored topic and to date the only systems

studied have been some diastereomeric salts[131]. This is despite the fact that pharmaceutical companies - one of the main proponents of developments in the field of crystal structure prediction - currently market an estimated 50 % of their compounds as organic salts[64]. Carboxylate salts in particular are the second most frequently occurring type of organic salt (excluding salts containing metal ions) stored in the Cambridge Structural Database (CSD) according to the results of a combined literature and CSD survey reported in Chapter 3 of this thesis. The unprecedented level of success reported in the 2007 blind test[212] of crystal structure prediction is encouraging and comes at a time when the focus is slowly beginning to shift to the crystal structures of multi-component solid forms. Future developments in the field will likely come from the advances made in methods used to estimate the relative lattice energies of hypothetical crystal structures with techniques based on electronic structure theory[139] already starting to show promise[216]. However, developments in the computing must run parallel with studies that use the wealth of structural information contained in the predicted crystal energy landscapes[217] to help address problems of practical importance to crystal engineers. It is with this in mind that we embark on the present study, the first to directly compare the predicted crystal energy landscapes of salts and cocrystals and in so doing, assess the degree to which such landscapes can be used to understand the crystallisation behaviour of salts and cocrystals.

5.2 Method

5.2.1 Experimental screen for a cocrystal of 4-cyanopyridine and 4-fluorobenzoic acid

Commercial samples of 4-cyanopyridine (Alfa Aesar, 98 % purity) and 4-fluorobenzoic acid (Sigma-Aldrich, 98 % purity) were used without further purification. Solution experiments were performed by lightly grinding (pestle and mortar) a stoichiometric 1:1 molar ratio of 4-cyanopyridine and 4-fluorobenzoic acid and dissolving the resulting powder in the minimum amount of solvent needed to dissolve 0.1 g of solute. The solvents screened were methanol, ethanol, acetone, tetrahydrofuran (THF) and water. In all cases, the solutions were filtered to remove any excess solute particles and the

solvent was allowed to evaporate at the temperatures: -4, 5 and 25 °C. Automated grinding experiments[54,187] were performed using a *Retsch MM200* mixer mill, equipped with 10 mL capacity stainless steel grinding jars and two 5 mm stainless steel grinding balls per jar. A stoichiometric 1:1 molar ratio of 4-cyanopyridine and 4-fluorobenzoic acid were used in neat grinding and solvent drop grinding experiments. In all cases, the combined mass of solute in the grinding jar did not exceed 0.5 g. Solvent drop grinding experiments were performed by adding 4 drops of solvent to the stoichiometric mixture of acid and base. Neat grinding experiments were performed at a frequency of 30 Hz for 60 minutes while solvent drop grinding experiments were performed at the same frequency but for 30 minutes. A PerkinElmer Spectra One Fourier Transform-Infrared (FT-IR) spectrometer was used to identify the ground samples. For each sample, spectra were collected at a resolution of 4 cm^{-1} using 15 scans in the wavenumber range 650-4000 cm^{-1}.

5.2.2 Single crystal X-ray diffraction

Single crystal X-ray diffraction experiments were performed on a Bruker AXS SMART APEX CCD diffractometer equipped with a Bruker AXS Kryoflex open flow cryostat [graphite monochromated Mo-K_α radiation (λ=0.71073 Å)]. Data integration and final unit cell parameters were obtained via SAINT+.[188] The absorption correction was performed via a semi-empirical approach using SADABS[189], and the crystal structure of **VIII** was solved by direct methods using SHELXS-97[190]. All atoms were located from the difference Fourier map. For the non-hydrogen atoms, an anisotropic model was used for the thermal parameters and the atomic co-ordinates freely refined. All hydrogen atom co-ordinates and isotropic thermal parameters were also freely refined. The SHELXL-97[191] package was used for structure refinement. Packing diagrams were produced using Mercury CSD 2.2.[176]

5.2.3 Computational modelling

The molecular structures of **V**, **VIII** and **IYUPEX** were extracted from the experimental crystal structures and used in separate gas phase *ab initio* geometry optimisations at the

MP2/6-31G (d,p) level of theory using GAUSSIAN03[119]. The *ab initio* geometry optimisations were repeated on the hypothetical molecular structures that result following proton transfer. For solid forms **V** and **VIII**, details on how to generate the hypothetical molecular structures can be found in section 4.2.4 of Chapter 4. For **IYUPEX**, both the salt and cocrystal positions for the acidic proton are observed and the desired atomic-connectivity was generated by removing the acidic proton from the N or O atoms. All editing of the molecular structures were done in Mercury CSD 2.2. In the *ab initio* geometry optimisation of the molecules and ions, the dihedral angles defining the methyl groups of 4-dimethylaminopyridine were constrained to the values found in the 4-dimethylaminopyridinium ion of **V** and the O-H bondlength of the maleate ion was constrained to the standard neutron[196] value of 1.015 Å. Both constraints were necessary to prevent the calculation of gas phase minima with features such as methyl pyramidalisation at nitrogen (4-dimethylaminopyridine) or an intramolecular proton that was roughly equidistant from the two oxygen atoms (maleate). Both these features are not found in the experimental or corresponding CSD entries of related multi-component crystal structures. The geometries of all other molecules/ions shown in Table 5.1 were optimised without any constraints.

In the first instance, crystal structure prediction for **V**, **VIII** and **IYUPEX** assumed the experimental atomic connectivity for the acidic proton but using the *ab initio* conformation of the molecules/ions. Then, the effects of proton transfer were investigated by applying the same crystal structure prediction methodology but using the atomic connectivity obtained following proton transfer from a salt to a cocrystal or vice versa. CrystalPredictor[88] was used to generate hypothetical crystal structures in the space groups $P\bar{1}$, $P2_1$, $P2_1/c$, $P2_12_12$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$, $Pbca$, $Pbcn$, Cc, $C2$, $C2/c$, $P2_1/m$ and $C2/m$. A total of 250,000 trial crystal structures were generated for each CrystalPredictor search involving the salt or cocrystal molecular pair of **V**, **VIII** and **IYUPEX**. Convergence was assessed on the basis of the number of times each of the low energy structures was found. For the CrystalPredictor searches of **V**, **VIII** and **IYUPEX** the global minimum structure was found more than five times, irrespective of whether the search employed the salt or cocrystal molecular structure. The dispersion-repulsion contributions towards the lattice energy were calculated using

an atom-atom potential model of the *exp*-6 type, with the parameters for C, H_C, O, N and F were taken from the work of Williams[102-104]. The parameters for the polar protons, H_N and H_O, were taken from the work of Coombes[105] and Beyer[106] respectively. In the CrystalPredictor searches, the electrostatic interactions were modelled using the atomic charges derived from fitting (CHELPG scheme[93]) to the electrostatic potential of the MP2/6-31G(d,p) *ab initio* wavefunction for the molecule/ion.

Following the CrystalPredictor searches, the 1000 lowest energy structures produced were passed to DMACRYS[89] for rigid body lattice energy minimisations using a more accurate electrostatic model obtained from a distributed multipole analysis (DMA) of the same MP2/6-31G(d,p) *ab initio* wavefunction for the isolated molecules/ions. The DMA was performed using the program GDMA2.2[195]. Only structures at true lattice energy minima were kept. These structures were clustered[193] to remove multiple findings of the same lattice energy minima. This was done by comparing the relative lattice energies and cell densities, simulated X-ray powder diffraction patterns and COMPACK[115] overlays of the 15-molecule co-ordination spheres ($RMSD_{15}$) using a similarity threshold of 0.1 Å or less for identical structures.

As in the previous chapter, the *Exp(Obs)MinOpt* and *Exp(Hyp)MinOpt* structures refer to the lattice energy minima calculated using *ab initio* optimised molecular/ionic conformations at the *observed* (Obs) or *hypothetical* (Hyp) acidic proton positions respectively. If one or more flexible degrees of freedom on the molecules/ions were constrained in the *ab initio* geometry optimisation, the resulting lattice energy minima are denoted *Exp(Obs)MinConOpt* or *Exp(Hyp)MinConOpt*.

The hydrogen bond motifs and associated graph sets of the predicted crystal structures were calculated with the Molecular Materials Studio feature of Mercury CSD 2.3[218]. To make the analysis of the hydrogen bond motifs feasible, only the predicted structures within 10 kJ mol^{-1} of the global minimum structure were analysed in the crystal energy landscapes of the salt and cocrystal. For solid forms **V** and **VIII**, the only intermolecular hydrogen bond motif observed was based on an acid-base/cation-anion interaction and the criterion used to distinguish between the $R_2^2(7)$ and $D_1^1(2)$ motifs was the same one used by Nangia[156]. This criterion states that when the C-O···N-C/C-O

$\cdots N^{+}$-C torsion angle is ≤ 15 º, the two molecules/ions are approximately coplanar and the motif is $R_2^2(7)$. If the torsion angle is > 15 º, the mean molecular planes of the hydrogen bonded molecules/ions are sufficiently different so as to imply the $D_1^1(2)$ motif. For **IYUPEX**, the graph sets used to identify the different hydrogen bond motifs refer to the acid-acid/anion-anion as well as the acid-base/cation-anion interactions. This is because there are a greater variety of motifs displayed by acid-acid/anion-anion interactions when compared to the motifs displayed by acid-base/cation-anion interactions.

For **V**, **VIII** and **IYUPEX**, the crystal packing effects of the acidic proton were investigated by re-positioning the acidic proton in each predicted salt structure within 10 kJ mol^{-1} of the salt energy landscape to give the cocrystal and relax the lattice energy of the resulting cocrystal. The structure editing was done in Mercury CSD 2.3 by generating the desired sp^3-hybridised O-H bond starting from the predicted salt lattice energy minimum. The same conformation and multipole model used in the cocrystal search was used to relax the lattice energy of the cocrystal that results from re-positioning the acidic proton. The packing similarity feature of Mercury CSD 2.3 was used to overlay all salt-cocrystal pairs using a 15 molecule co-ordination sphere ($RMSD_{15}$) and thresholds of 30 % for the intermolecular distances and 30° in the angles. The calculated $RMSD_{15}$ depends on the positions of only the non-hydrogen atoms and this allows us to compare the crystal structures of salts and cocrystals. Visual inspection of the acid-base/cation-anion hydrogen bond motif was also performed for all salt-cocrystal pairs. The effects of energetic re-ranking were quantified by comparing the energy of the salt and cocrystal alternatives relative to the predicted global minimum in each landscape.

5.3 Crystal structure of 4-cyanopyridine 4-fluorobenzoic acid (1:1)

Solution crystallisation of a 1:1 molar ratio of 4-cyanopyridine and 4-fluorobenzoic acid in methanol, ethanol, acetone and tetrahydrofuran (THF) all led to block-like crystals of a 4-cyanopyridine 4-fluorobenzoic acid cocrystal, **VIII**, with a 1:1 molar ratio of the molecules (Figure 5.2). Both coformers were found to be insoluble in water and as a

4-cyanopyridine & 4-fluorobenzoic acid (1:1 molar ratio)				
Solvent	**Solvent Evaporation (25 °C)**	**Solvent Evaporation (5 °C)**	**Solvent Evaporation (-4 °C)**	**Solvent Drop Grinding (30 mins, frequency 30 Hz)**
Methanol	**VIII**	**VIII**	**VIII**	**VIII**
Ethanol	**VIII**	**VIII**	**VIII**	**VIII**
Acetone	**VIII**	**VIII**	**VIII**	**VIII**
THF	**VIII**	**VIII**	**VIII**	**VIII**
Water*	-	-	-	-

Table 5.2: The above table shows the experimental matrix of solid forms obtained following cocrystallisation of a 1:1 molar ratio of 4-cyanopyridine and 4-fluorobenzoic acid. **VIII**= 4-cyanopyridine 4-fluorobenzoic acid (1:1) cocrystal. Identification was by IR. * Both the coformers were insoluble in water.

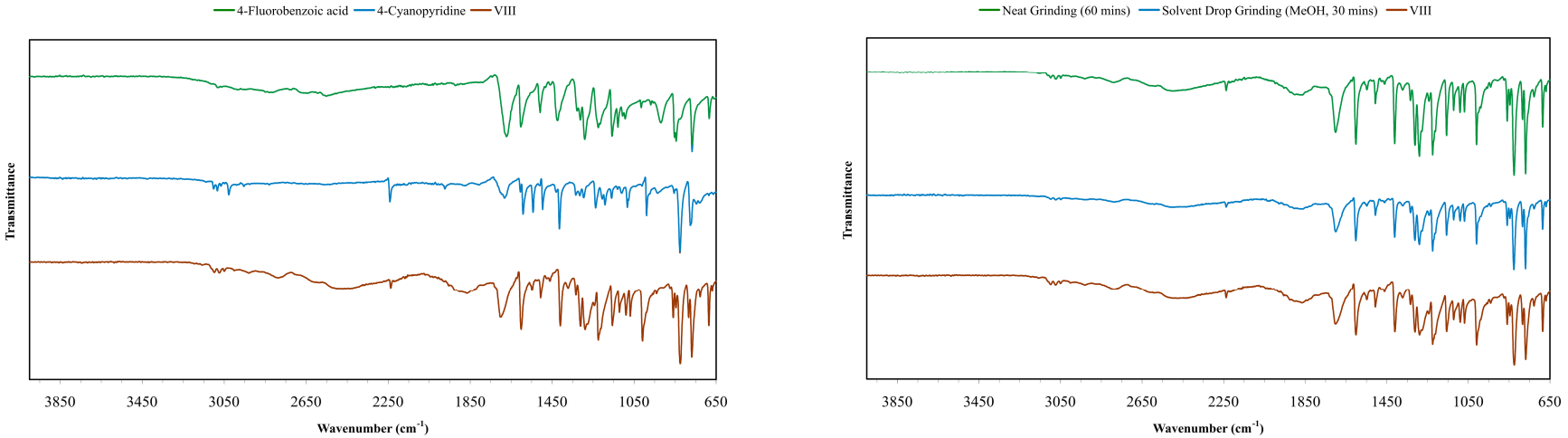


Figure 5.1: Shown in a) is a comparison of the characteristic FT-IR spectra of the 4-cyanopyridine and 4-fluorobenzoic acid starting reagents with the spectrum of **VIII** while b) shows that the results of neat grinding and solvent drop grinding (Methanol, MeOH) produce the cocrystal, **VIII**, obtained from the solution crystallisation experiments. Solvent drop grinding with ethanol, acetone and THF produced the same spectrum shown above for methanol.

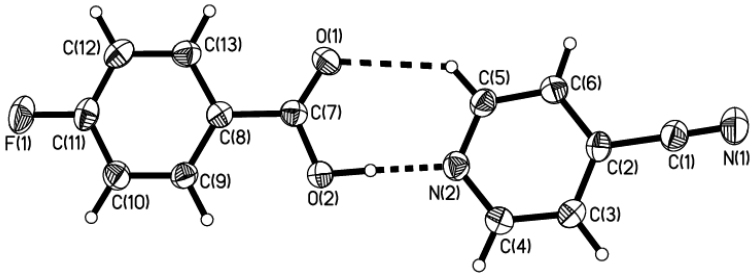


Figure 5.2: The asymmetric unit of the 4-cyanopyridine 4-fluorobenzoic acid (1:1) cocrystal, **VIII**. Displacement ellipsoids are drawn at the 50 % probability level and hydrogen atoms are shown as spheres of arbitrary radii.

	VIII
Formula	$C_{13}H_9FN_2O_2$
Crystal system	Monoclinic
Space group	C2/c
***a* (Å)**	12.959(3)
***b* (Å)**	7.4852(15)
***c* (Å)**	23.703(5)
a (°)	90
b (°)	91.632(3)
g (°)	90
V (Å^3)	2298.3(8)
Z	8
T (K)	150
F(000)	1008
D_{calc} (g cm^{-3})	1.412
μ (mm^{-1})	0.108
Crystal size (mm^3)	$0.55 \times 0.30 \times 0.24$
Reflns collected	9124
Unique reflns (R_{int})	2685 (0.0232)
GOOF on F^2	1.059
R_1 [$F^2 > 2s(F^2)$]	0.0401
wR_2 (all data)	0.1141
Largest difference map features (e Å^{-3})	0.260, -0.203

Table 5.3: Crystallographic parameters for the novel 4-cyanopyridine 4-fluorobenzoic acid cocrystal, **VIII**.

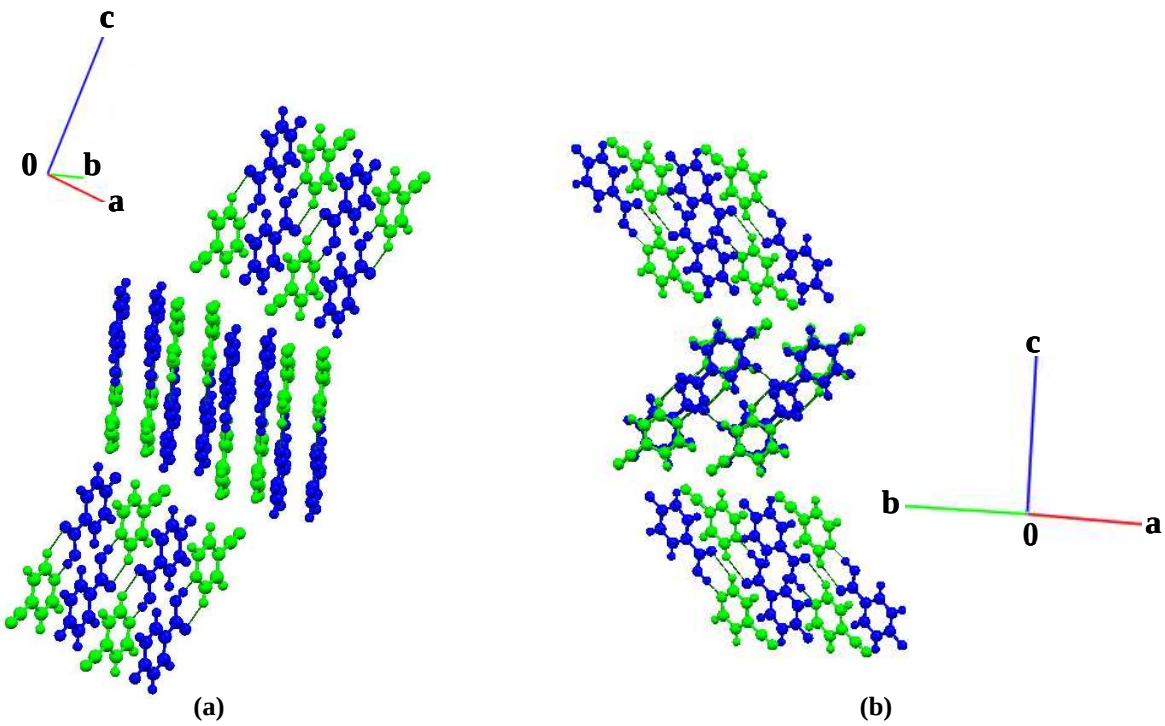


Figure 5.3: Part (a) illustrates the molecular heterodimers between 4-fluorobenzoic acid (blue) and 4-cyanopyridine (green) in the crystal structure of **VIII**. The packing is such that the heterodimers do not form two-dimensional sheets. Part (b) is an alternative view of the dimer packing and emphasises the π-π stacking (centroid-centroid separation of 3.78 Å) between acid and base molecules of inversion related dimers for the middle layer.

consequence no solution or grinding experiments were performed that employ this solvent. The results of neat grinding and solvent drop grinding experiments were identified by comparing the FT-IR spectra of these samples with the spectrum of **VIII**. In all cases, the IR spectrum of the sample obtained following solvent drop grinding was the same as that following neat grinding and both matched the IR spectrum of **VIII** (Figure 5.1). Characteristic vibrational bands that led to this identification include the nitrile stretching vibration which has a characteristic absorption in the wavenumber range 2200-2300 cm^{-1}. A summary of the experimental screen for a cocrystal of 4-cyanopyridine and 4-fluorobenzoic acid is given in Table 5.2 and Table 5.3 gives the crystallographic parameters for the cocrystal, **VIII**. Figure 5.1 shows the characteristic IR spectra of the pure components, powder samples obtained following grinding experiments and **VIII**.

The 4-cyanopyridine 4-fluorobenzoic acid cocrystal, **VIII**, crystallises in the monoclinic space group $C2/c$. The hydrogen bond motif between the acid and base is

characterised by a heterodimer of graph set $R_2^2(7)$ and is possible because the 4-cyanopyridine and 4-fluorobenzoic acid molecules are approximately coplanar with an angle of 3.65 ° between the mean molecular planes of the acid and base. Other intermolecular interactions include the C-H···N≡C contact between two molecules of 4-cyanopyridine as well as a C-H···F-C interaction between 4-cyanopyridine and 4-fluorobenzoic acid. Both interactions are facilitated by the same C-H group on the base with respective intermolecular distances of 2.703 Å and 2.561 Å between the hydrogen atom and acceptor groups. Figure 5.3 illustrates the packing of the dimers in the crystal structure. The same molecules in the heterodimer are not stacked on top of another as a consequence of the inversion operator in the crystal. However the acid and base do display π-π stacking for the aromatic rings (part b of Figure 5.3).

5.4 Crystal energy landscapes for 1:1 salts and cocrystals

5.4.1 4-dimethylaminopyridinium maleate (1:1) salt

The calculated crystal energy landscape of 4-dimethylaminopyridinium maleate, **V**, is shown in Figure 5.4. The experimental structure of **V** is predicted to be the most stable in the crystal energy landscape of the salt (Figure 5.4) and there is a 2 kJ mol^{-1} energy gap between the global minimum structure and the next most stable structure in the calculated crystal energy landscape. The finding that the experimental and predicted most stable structures of 4-dimethylaminopyridinium maleate are the same represents a success in the crystal structure prediction of carboxylate salts. This success is highlighted in Figure 5.6 which shows a match following an overlay of the crystal packings of **V** and the predicted global minimum structure from the search. The root mean square deviation for the overlay of the 15-molecule co-ordination spheres ($RMSD_{15}$) of the two structures is 0.153 Å.

The crystal energy landscape of the hypothetical 4-dimethylaminopyridine maleic acid cocrystal is shown in Figure 5.5. In the crystal energy landscape of 4-dimethylaminopyridine maleic acid (Figure 5.5), the nearest matching hypothetical cocrystal structure is ranked 18th in stability and is approximately 2.41 kJ mol^{-1} less stable than the predicted global minimum structure. The $RMSD_{15}$ for the overlay of **V**

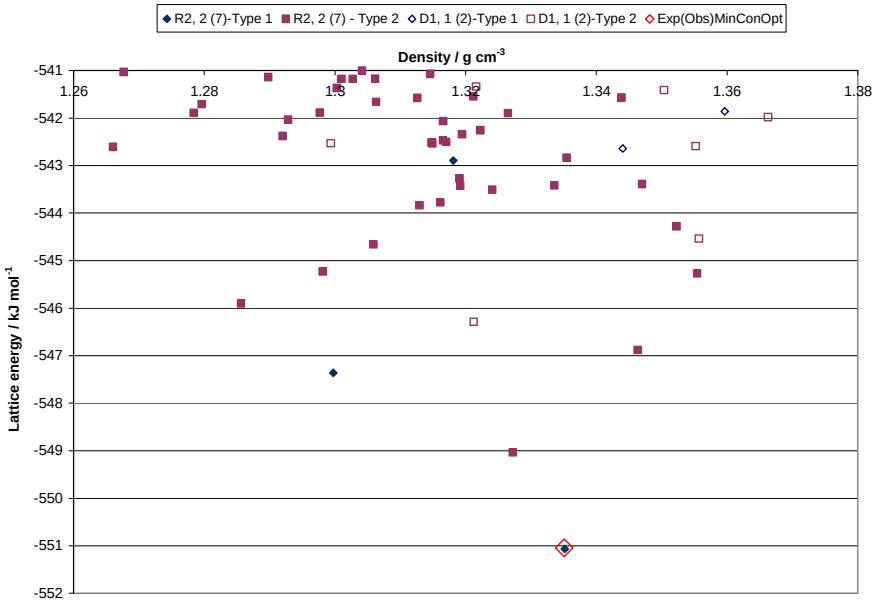


Figure 5.4: Scatter plot of the lattice energy versus density for the predicted crystal structures of 4-dimethylaminopyridinium maleate, **V**. Only the lattice energy minima within 10 kJ mol^{-1} of the global minimum structure are shown and all structures are classified according to the graph set of the hydrogen bond motif. *Exp(Obs)MinConOpt* denotes the experimental structure of **V**.

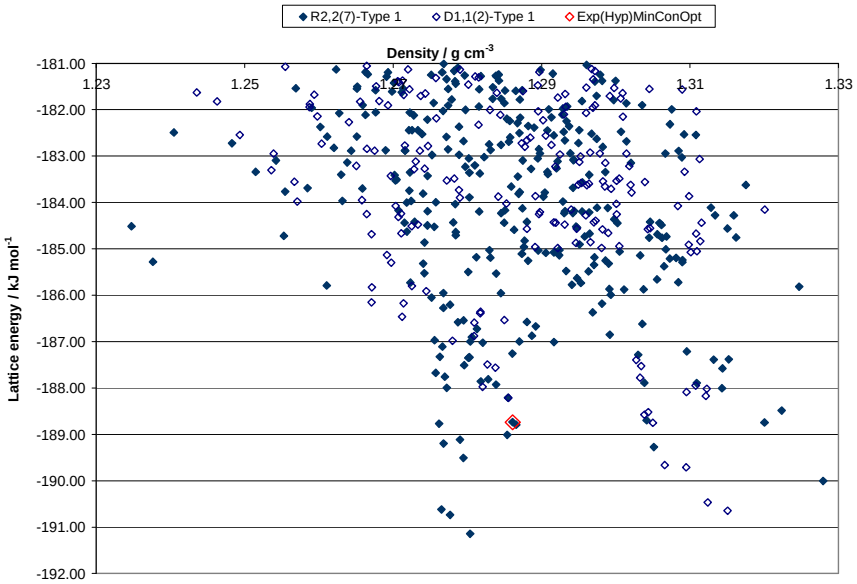


Figure 5.5: Scatter plot of the lattice energy versus density for the predicted crystal structures of the hypothetical 4-dimethylaminopyridine maleic cocrystal. *Exp(Hyp)MinConOpt* is derived from **V** following intermolecular proton transfer and subsequent lattice energy minimisation assuming the *ab initio* optimised molecular conformations for maleic acid and 4-dimethylaminopyridine.

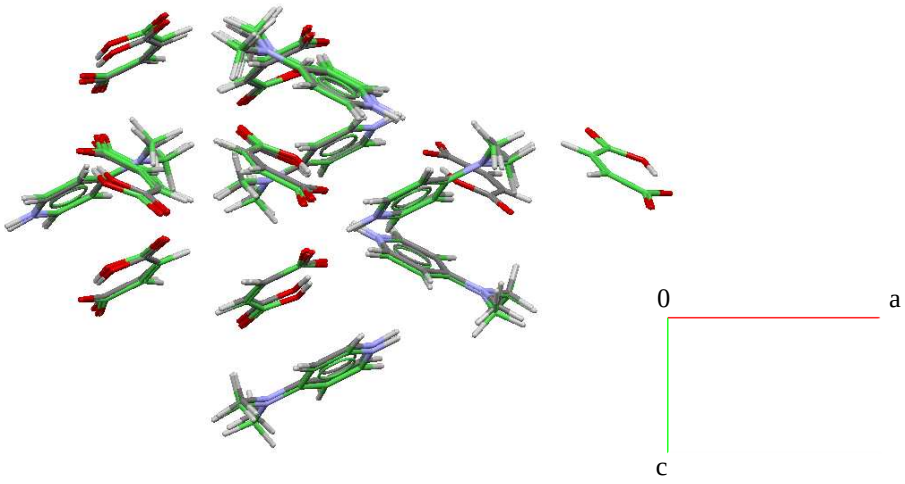


Figure 5.6: Overlay of the crystal packings for **V** (coloured by element) and the predicted global minimum structure (coloured green) from the salt search.

and the lattice energy minimum of the hypothetical cocrystal is 0.488 Å indicating a poorer match in the crystal packings when an incorrect acidic proton position is assumed in the modelling. Within the same energy range of 10 kJ mol^{-1} from the global minimum structure, the salt crystal energy landscape is much less populated than the cocrystal energy landscape. However, if we plot the lattice energies of the predicted structures as a % of the global minimum lattice energy, we find that within 4 % of the total lattice energy for the global minimum structure, there are 261 predicted structures in the cocrystal energy landscape and 433 structures in the salt energy landscape. This indicates the difficulties in comparing the crystal energy landscapes of salts and cocrystals.

Inspection of the hydrogen bond motifs among the low energy structures shows that in both crystal energy landscapes, the lattice energy minima adopt the $R_2^2(7)$ or $D_1^1(2)$ graph sets. Visual inspection of the hydrogen bond motifs in the salt and cocrystal energy landscapes has shown two versions of the $R_2^2(7)$ and $D_1^1(2)$ motifs differing in the oxygen atom used for the N^+-H···O hydrogen bond. These are illustrated in Table 5.4 and are denoted Type 1 or Type 2. The electron density in the carboxylate group is not equally distributed among the two oxygen atoms as a consequence of the intramolecular hydrogen bond in the maleate ion. Thus in the

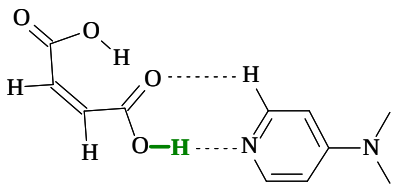
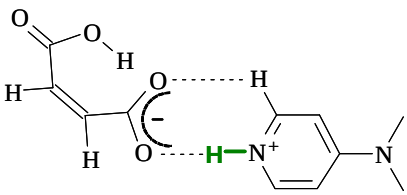
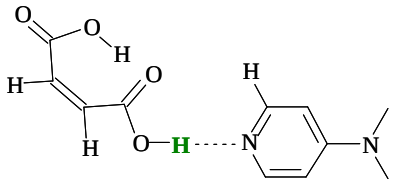
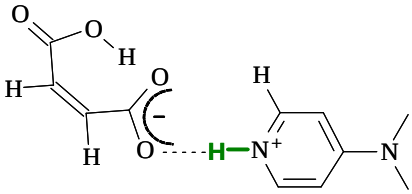
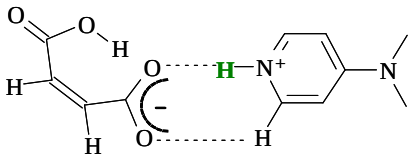
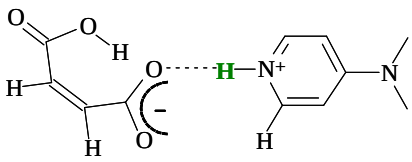
Hydrogen bond motifs observed in cocrystal energy landscape	Hydrogen bond motifs observed in salt energy landscape
 R2,2(7)-Type 1	 R2,2(7)-Type 1
 D1,1(2)-Type 1	 D1,1(2)-Type 1
	 R2,2(7)-Type 2
	 D1,1(2)-Type 2

Table 5.4: Illustration of the hydrogen bond motifs observed in the salt and cocrystal energy landscapes of **V.** The acidic proton used for intermolecular hydrogen bonded is highlighted in green. The predicted salt structures display a greater variety of hydrogen bond motifs owing to the combined effects of the intramolecular hydrogen bond and the delocalisation of charge in the carboxylate group.

crystal energy landscape of the salt, there are two types of $R_2^2(7)$ and $D_1^1(2)$ motifs (Table 5.4), which are distinguished by the oxygen atom used for hydrogen bonding with the N^+-H donor of the cation. Type 2 interactions are more prevalent for salt structures with the heterodimer $R_2^2(7)$ graph set as well as those with the single point

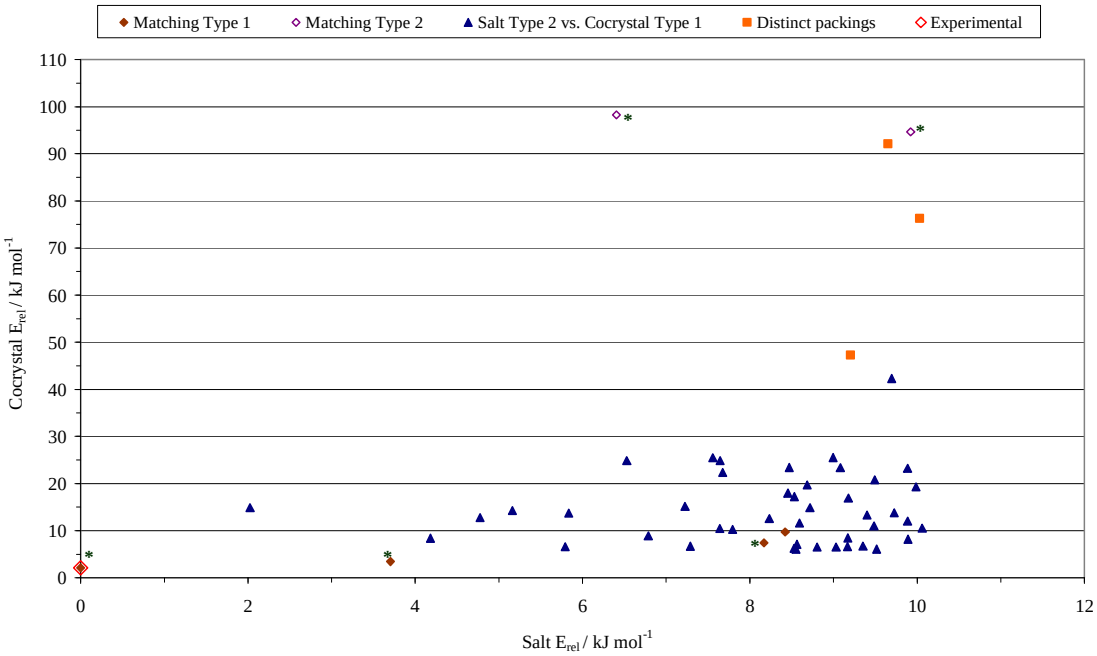


Figure 5.7: Packing comparisons of the 53 salt structures within 10 kJ mol^{-1} of global minimum structure of **V** and their nearest cocrystal counterparts obtained following proton transfer. The salt and cocrystal energies are relative to the global minimum structure in each landscape. * These are comparisons where the salt and matching cocrystal structures have sufficiently similar gross packings so as to overlay all 15/15 molecules in the co-ordination spheres.

$D_1^1(2)$ graph set. In fact, 41/44 salt structure with the $R_2^2(7)$ dimer motif are of Type 2 and 7/9 salt structures with the $D_1^1(2)$ motif are also of Type 2. Notably, the experimental salt structure of **V**, which is predicted as the global minimum in Figure 5.4, adopts the less common $R_2^2(7)$-Type 1 motif. In the crystal energy landscape of the cocrystal, only the Type 1 $R_2^2(7)$ and $D_1^1(2)$ motifs are observed within 10 kJ mol^{-1} of the global minimum structure (Figure 5.5). This is consistent with physical expectation as only the Type 1 interactions lead to the matching of complementary regions of positive (O-H) and negative (N) potentials on the acid and base.

Computational experiments to relocate the acidic protons in the 53 salt structures shown in Figure 5.4 have revealed that only 5 salt structures are matches with their nearest alternative cocrystal structures using the $RMSD_{15}$ (30% distance, 30 ° angle tolerance) packing similarity feature of Mercury. The alternative cocrystal structures were calculated by relocating the acidic proton to give the acid-base pair within the crystal and relaxing the lattice energy using the same conformation and multipole model

used in the cocrystal search (Figure 5.5). The results from these calculations are shown in Figure 5.7, which plots the relative energy of each salt against the relative energy of the matching cocrystal. Of the five structures that matched 15/15 molecules in $RMSD_{15}$, 3 had the $R_2^2(7)$-Type 1 motif in both the salt and matching cocrystal, and the experimentally observed salt structure was among these. Surprisingly the remaining two matches were salt structures that maintained the $R_2^2(7)$-Type 2 motif in the cocrystal. The resulting cocrystals have relative energies of 94.66 kJ mol^{-1} and 92.26 kJ mol^{-1} making them highly unlikely to be observed under any solution crystallisation conditions that may modify the ΔpK_a of the acid-base pair. The majority of the comparisons failed to match all 15/15 molecules in packing similarity because proton transfer induced a change from the Type 2 $R_2^2(7)$ or $D_1^1(2)$ motif to the preferred Type 1 in the cocrystal.

In the crystal energy landscape of the salt, structures with the $R_2^2(7)$ motif [Type 1 or Type 2] are energetically favoured over those with the $D_1^1(2)$ motif. Indeed the first structure that bears the $D_1^1(2)$-Type 1 motif in the crystal energy landscape of the salt has a relative lattice energy of 8.40 kJ mol^{-1} with respect to the experimental structure. In the crystal energy landscape of the cocrystal, the reverse is true and structures based on the $D_1^1(2)$-Type 1 motif are among the most stable structures. This indicates that an incorrect assignment of the acidic proton position can lead to significant re-ranking in the energetic stability of structures based on the same hydrogen bond motifs.

5.4.2 Pyridine isophthalic acid : pyridinium isophthalate (58%:42%)

The calculated crystal energy landscape of the 1:1 pyridine isophthalic acid cocrystal is shown in Figure 5.8 and Figure 5.9 shows the crystal energy landscape of the 1:1 pyridinium isophthalate salt. The molecular structures of the cocrystal and salt (Table 5.1) are observed in **IYUPEX** according to the percentage ratio of 58%:42% respectively. In the calculated crystal energy landscape of the cocrystal (Figure 5.8), the cocrystal major product is ranked second in stability and is only 0.16 kJ mol^{-1} less stable

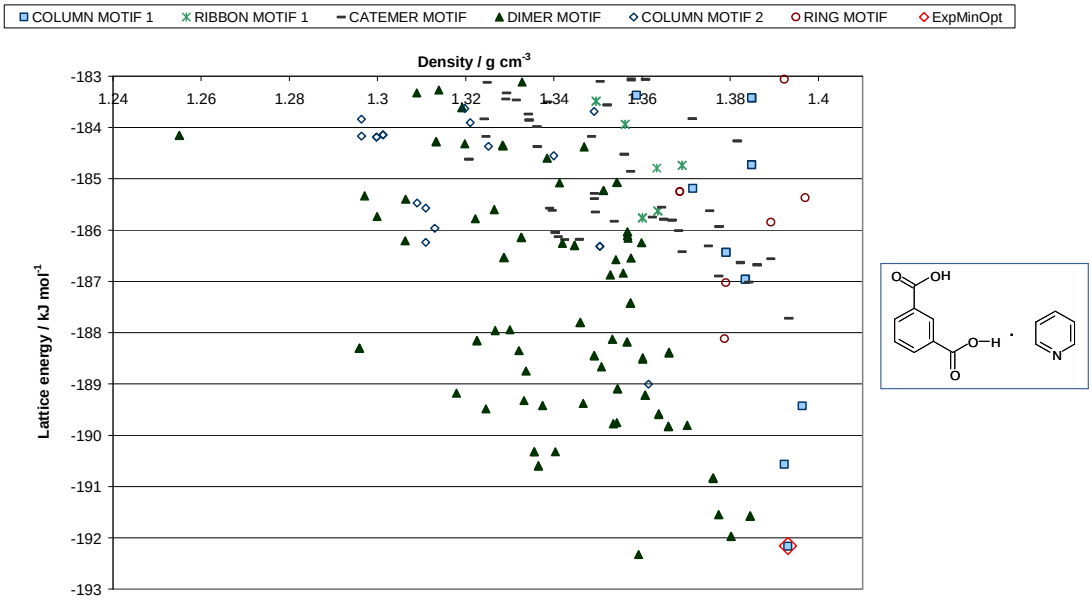


Figure 5.8: Scatter plot of the lattice energy versus density for the predicted crystal structures of pyridine isophthalic (1:1) cocrystal, **IYUPEX**. The cocrystal molecular diagram shown to the right of the scatter plot is the major product from the crystallographic refinement of the acidic proton in **IYUPEX**.

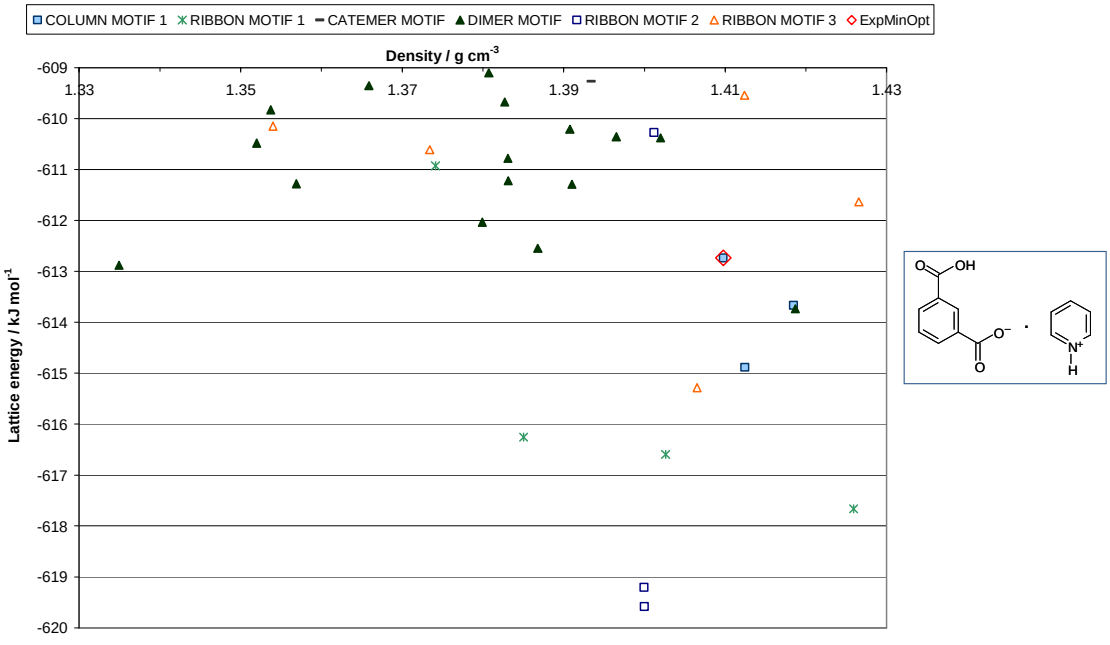
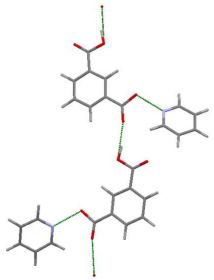
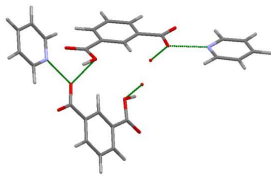


Figure 5.9: Scatter plot of the lattice energy versus density for the predicted crystal structures of the pyridinium isophthalate (1:1) salt, which appeared as the minor product from the crystallographic refinement of the acidic proton position in **IYUPEX**.



COLUMN MOTIF 1

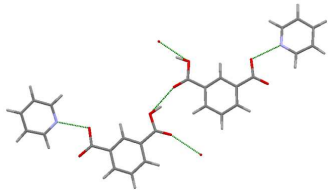
$C_1^1(8)$



RIBBON MOTIF 1

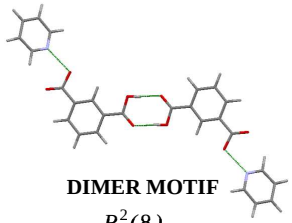
$C_1^1(8)$

HETERO MOTIF:	$R_2^2(7)$	$D_1^1(2)$	$R_2^2(7)$	$D_1^1(2)$
COO⁻···H-N⁺(CH) (Salt)	***0***	***3***	***4***	***0***
COO-H ···N(CH) (Cocrystal)	***4***	***9***	***9***	***2***



CATEMER MOTIF

$C_1^1(4)$

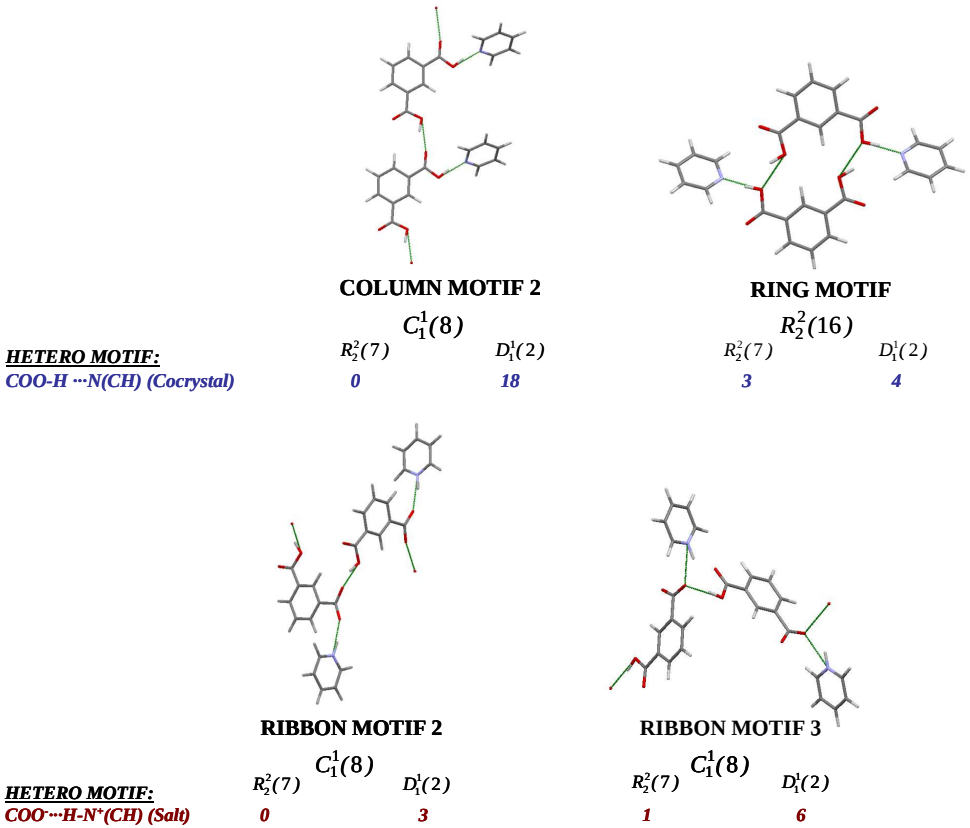


DIMER MOTIF

$R_2^2(8)$

HETERO MOTIF:	$R_2^2(7)$	$D_1^1(2)$	$R_2^2(7)$	$D_1^1(2)$
COO⁻···H-N⁺(CH) (Salt)	***0***	***1***	***5***	***11***
COO-H ···N(CH) (Cocrystal)	***39***	***22***	***51***	***42***

Scheme 5.1: Illustration of the acid-acid/anion-anion hydrogen bond motifs found in the crystal energy landscapes of the salt and cocrystal molecular structures of **IYUPEX**. The first graph set under the motif name refers to the illustrated acid-acid interaction. This is further subdivided into the number of structures with the $R_2^2(7)$ or $D_1^1(2)$ hetero motif between acid-base/cation-anion. The location of the ionisable proton is removed from all illustrations so as to apply to the salt/cocrystal.



Scheme 5.2: Illustration of the acid-acid hydrogen bond motifs that are unique to the salt or cocrystal energy landscapes of **IYUPEX**. The first graph set under the motif name refers to the illustrated acid-acid interaction. A further sub-classification of the number of structures with the $R_2^2(7)$ or $D_1^1(2)$ hetero motif between acid-base/cation-anion is provided.

than the global minimum structure. The cluster of low energy minima indicates that there is no single preferred structure on the basis of lattice energy. An overlay of the structures of **IYUPEX** and the matching predicted lattice energy minimum from the cocrystal search led to a calculated $RMSD_{15}$ of 0.239 Å.

Crystal structure prediction of the pyridinium isophthalate salt of **IYUPEX** has shown that the salt minor product is ranked 11th in stability and is 6.85 kJ mol^{-1} less stable than the global minimum structure of the salt search (Figure 5.9). An overlay of the structures of **IYUPEX** and the matching lattice energy minimum from the salt search led to a calculated $RMSD_{15}$ of 0.475 Å. Comparing this with the previously calculated $RMSD_{15}$ for overlaying **IYUPEX** and the predicted cocrystal structure indicates that the cocrystal is a better match to the experimental **IYUPEX** structure when compared to the predicted salt structure.

Scheme 5.1 illustrates the hydrogen bond motifs common to the crystal energy landscapes of the salt and cocrystal molecular structures of **IYUPEX**. Scheme 5.2 illustrates the unique hydrogen bond motifs that are found in the crystal energy landscape of the salt but not the cocrystal and vice versa. The expected O-H···N hydrogen bond interaction is found in all the predicted structures of the cocrystal and the resulting motif displays either the $R_2^2(7)$ or $D_1^1(2)$ graph set between acid and base. The acid-base $R_2^2(7)$ motif occurs in 52 % of cocrystal structures. By contrast, the N^+-H···O hydrogen bond in the salt can involve either of the two carboxylate oxygen atoms, but there are in total 10/34 salt structures with the $R_2^2(7)$ dimer motif between the ions.

According to Scheme 5.1, the majority of the predicted structures in the salt and cocrystal energy landscapes display the $R_2^2(8)$ carboxylic acid homodimer motif. The $R_2^2(8)$ carboxylic acid homodimer occurs in greater than 45 % of structures in the crystal energy landscapes of the salt and cocrystal. This makes it the most frequently occurring acid-acid interaction in the predicted salt and cocrystal structures of **IYUPEX**. There appears to be a preference for the $D_1^1(2)$ motif between cation-anion when the salt displays the $R_2^2(8)$ carboxylic acid homodimer motif. In the predicted cocrystal structures with the $R_2^2(8)$ carboxylic acid homodimer motif, there appears to

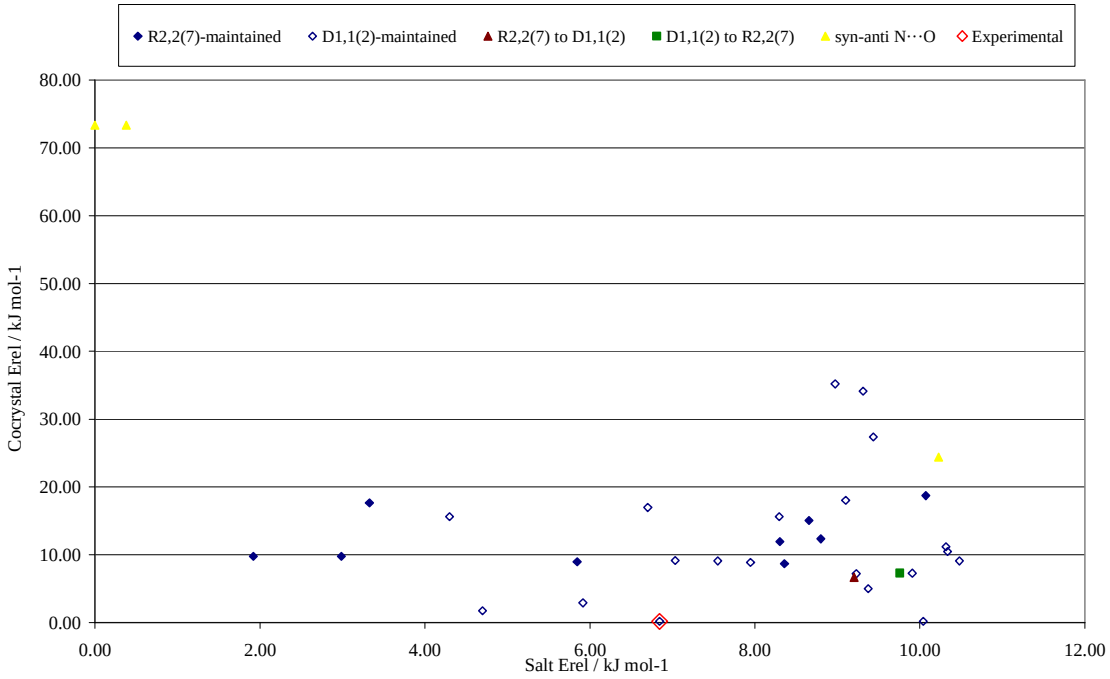


Figure 5.10: Changes in the cation-anion motif of the 32 salt structures in the predicted crystal energy landscape of **IYUPEX** (Figure 5.9) following simulations that relocate the acidic proton from a salt to a cocrystal. The salt and cocrystal energies are relative to the global minimum structure in each landscape.

be a roughly even split in the number of structures with the $R_2^2(7)$ or $D_1^1(2)$ motif between acid-base. If we compare the crystal energy landscapes shown in Figure 5.8 and Figure 5.9, we find that structures based on the $R_2^2(8)$ carboxylic acid homodimer are energetically favoured in the cocrystal energy landscape but unfavoured in the salt energy landscape. Such re-ranking in the stabilities of structures based on the same motifs was observed in the $D_1^1(2)$-Type 1 motif of **V** (*cf.* Figure 5.4 and Figure 5.5). As seen in **V**, a further effect of proton transfer in **IYUPEX** is the finding of hydrogen bond motifs that are exclusive to the salt or cocrystal energy landscape. For example, the two most stable structures in the salt energy landscape are based on Ribbon Motif 2 (Scheme 5.2), which is not observed in the cocrystal energy landscape. Two of the three salt structures that display Ribbon Motif 2, use the anti-lone pair on the carboxylate oxygen (illustrated in Scheme 5.2) for the N^+-H···O hydrogen bond interaction. The structures concerned are the two most stable lattice energy minima in the salt energy landscape (Figure 5.9).

Computational experiments that relocate the acidic proton (salt→cocrystal) in the 34 salt structures of Figure 5.9 showed that relaxing the alternative cocrystal structure led to a change in the oxygen used for the N^+-H···O intermolecular interaction in 13 out of 34 structures, but the experimentally observed salt minor product was not among these. In all cases, changes in the atoms used for the cation-anion/acid-base interaction did not induce significant changes in the packing of the anion/acid molecules within the crystal. Figure 5.10 shows the changes in the relative energies of the salt and cocrystal structures as well as the changes in the acid-base/cation-anion motif of **IYUPEX** in simulations that go from the salt to the cocrystal molecular structure. The majority of salt structures (20/34) maintained the $D_1^1(2)$ motif between cation-anion to give the cocrystal equivalent between the acid-base pair. There were only two examples where proton transfer was sufficient to cause a change from $R_2^2(7)$ to $D_1^1(2)$ or vice versa. Both the global minimum and second ranked salt structures are unstable as cocrystals and this is reflected in the relative energy of 73.35 kJ mol^{-1} (Figure 5.10) for the hypothetical cocrystals. Both these structures maintain use of the oxygen anti-lone pair for the cocrystal N···O interaction but the intermdecular distances of 3.20 Å are too long to be defined as hydrogen bonds in Mercury. This contrasts with the N^+···O intermolecular distances in the salt versions of these structures, which do not exceed 2.69 Å. Overall, it appears that the competing acid-acid and acid-base/cation-anion hydrogen bonds in **IYUPEX** leads to greater complexity in the possible predicted packings (Figure 5.8 and Figure 5.9). In terms of the crystal packing effects of acidic proton position, 19/34 salt-cocrystal pairs matched 15/15 molecules in $RMSD_{15}$.

Direct comparison of the lattice energies of the predicted salt and cocrystal structures of **IYUPEX** is not possible because the electrostatic contribution towards the lattice energies of salts is much greater than that of cocrystals. This means that salts have much more stable lattice energies than cocrystals and this is why the energy scales of Figure 5.8 and Figure 5.9 differ by more than a factor of 3. From the electronic structure modelling of **IYUPEX** presented in the previous chapter (section 4.4.3), an estimate of the energy penalty to relocate the acidic proton from the salt to the cocrystal (Figure 4.14) has shown that the most stable molecular structure has an N^+-H bondlength of 1.15 Å. Such a bondlength is atypical of both salts and cocrystals and is

in line with the observed crystallographic disorder in the acidic proton position. The most stable cocrystal molecular structure has a relative lattice energy of 7.43 kJ mol^{-1}.

5.4.3 4-cyanopyridine 4-fluorobenzoic acid (1:1) cocrystal

The calculated crystal energy landscape of the 4-cyanopyridine 4-fluorobenzoic acid cocrystal, **VIII**, is shown in Figure 5.11. The experimental 4-cyanopyridine 4-fluorobenzoic acid cocrystal is among the most stable of the predicted crystal structures in Figure 5.11 and is approximately 1.87 kJ mol^{-1} less stable than the predicted global minimum structure. An overlay of the crystal packings of **VIII** and the matching predicted lattice energy minimum from the search led to a calculated $RMSD_{15}$ of 0.466 Å. This is a poorer match than was observed for similar overlays involving the experimental structure of **V** and the disordered major product of **IYUPEX**.

The crystal energy landscape of the hypothetical 4-cyanopyridinium 4-fluorobenzoate salt is shown in Figure 5.12. The hypothetical salt structure is predicted to be 8.31 kJ mol^{-1} less stable than the predicted global minimum in the salt crystal energy landscape. Thus if we compare the salt and cocrystal energy landscapes on a qualitative basis, the correct molecular structure is calculated to give a crystal among the most stable of the predicted structures while the incorrect molecular structure is not. An overlay of the crystal packings of **VIII** and the nearest matching lattice energy minimum from the salt search led to an $RMSD_{15}$ of 0.416 Å.

Inspection of the hydrogen bond motifs in the predicted crystal structures of Figure 5.11 and Figure 5.12 show that the possible hydrogen bond motifs are based on the $R_2^2(7)$ or $D_1^1(2)$ graph sets. For the predicted salt structures of **VIII** (Figure 5.12), the molecular C_{2V} symmetry of both ions means that the carboxylate oxygens on the anion are symmetrically equivalent and hence different versions of the $R_2^2(7)$ and $D_1^1(2)$ graph sets are not seen (contrast with **V**). For the predicted cocrystal structures of **VIII** (Figure 5.11), the expected O-H···N interaction isseen for all structures within 10 kJ mol^{-1}, and the structures differ in the coplanarity of the acid-base pair [*i.e.* $R_2^2(7)$ and $D_1^1(2)$]. In the cocrystal energy landscape (Figure 5.11) of **VIII**, 55 % of

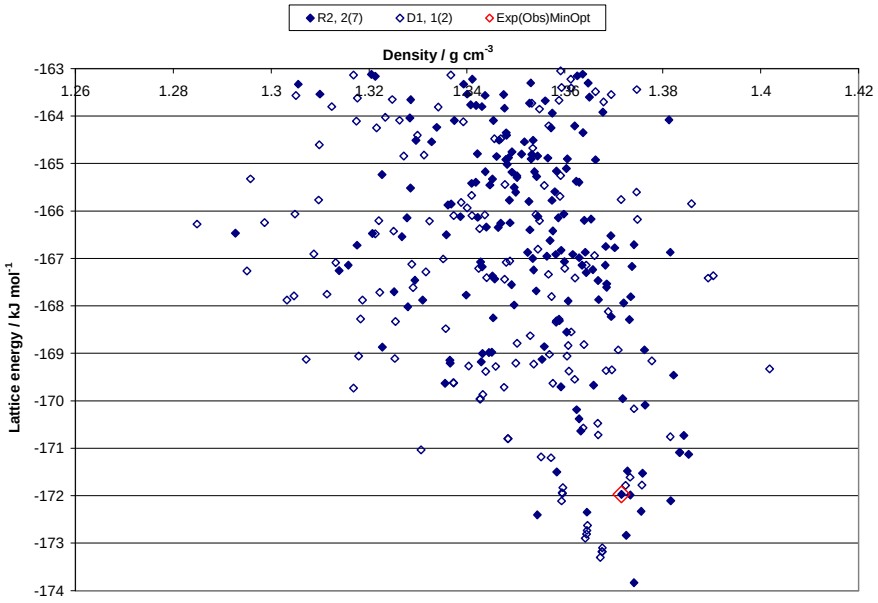


Figure 5.11: Scatter plot of the lattice energy versus density for the predicted crystal structures of the 4-cyanopyridine 4-fluorobenzoic acid cocrystal, **VIII**. Only the lattice energy minima within 10 kJ mol^{-1} are shown and the structures are classified according to the graph set of the hydrogen bond motif. *Exp(Obs)MinOpt* denotes the experimental structure of **VIII**.

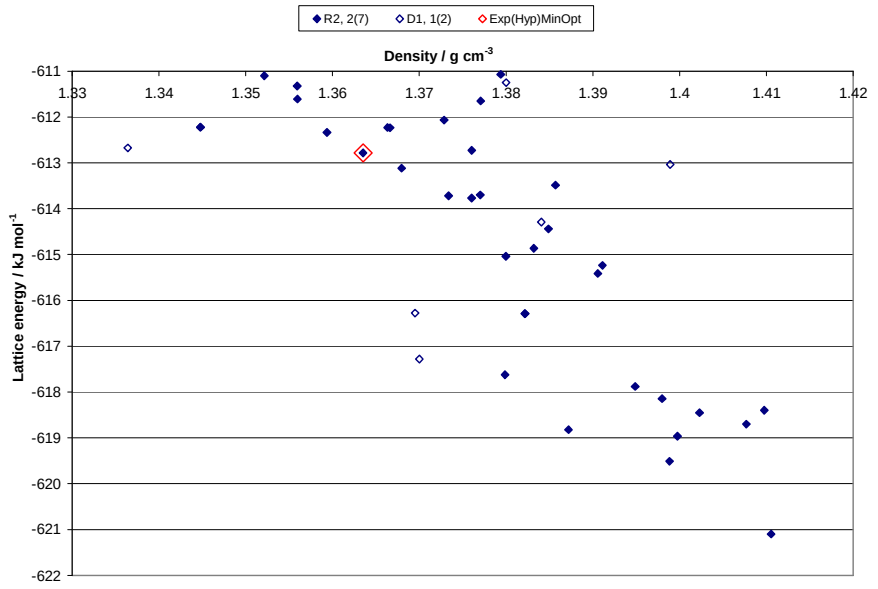


Figure 5.12: Scatter plot of the lattice energy versus density for the predicted crystal structures of the hypothetical 4-cyanopyridinium 4-fluorobenzoate salt. *Exp(Hyp)MinOpt* is derived from **VIII** following intermolecular proton transfer and subsequent lattice energy minimisation assuming the *ab initio* optimised molecular conformations of the cation and anion.

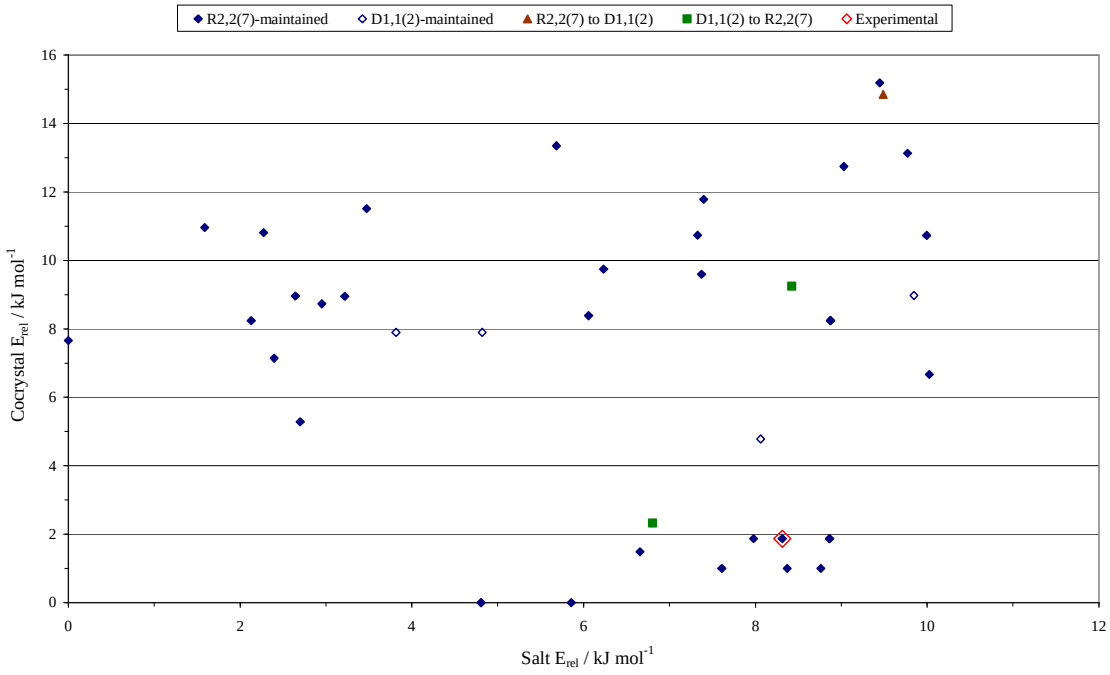


Figure 5.13: Changes in the packing and relative energies of the 41 hypothetical salt structures of **VIII** in simulations that relocate the acidic proton from the salt to the cocrystal. The true molecular structure of **VIII** is that of a cocrystal. The energies of the salt and cocrystal structures are relative to the global minimum structure in the respective crystal energy landscape.

structures display the $R_2^2(7)$ dimer motif between acid and base, while in the salt (Figure 5.12) crystal energy landscape the same motif accounts for 85 % of structures.

As in **V** and **IYUPEX**, computational experiments to relocate the acidic proton from the salt to the cocrystal were performed to investigate the structural changes that occur within the predicted crystals. The results from these calculations are illustrated in Figure 5.13, which plots the relative energy of the salt and nearest matching cocrystal. Each salt-cocrystal pair is denoted in terms of the observed changes in the packing motif. The calculations revealed that 38/41 structures in the salt energy landscape maintained the $R_2^2(7)$ or $D_1^1(2)$ motif in the cocrystal, and 32 of the 38 salt structures that maintained the motif had a cocrystal alternative with 15/15 molecules matched in $RMSD_{15}$. The experimental structure of **VIII** was among the salt-cocrystal pairs that maintained the $R_2^2(7)$ motif upon proton transfer and matched 15/15 molecules in packing similarity. The fact that the majority of cocrystal structures maintained the $R_2^2(7)$ or $D_1^1(2)$ motif observed in the hypothetical salt suggests that proton transfer

does not have the same dramatic effect in creating new hydrogen bond motifs as observed for **V** and **IYUPEX**.

Although the experimental cocrystal structure of **VIII** is among the most stable structures (Figure 5.11), there are a handful of alternative low energy structures that are more stable than it. By contrast, the crystal energy landscapes of **V** and **IYUPEX** show that the experimental structure or disordered major product is either the global minimum or energetically competitive with the global minimum in lattice energy. The dispersion-repulsion model used is clearly not as suitable for **VIII** as it is for **V** and **IYUPEX**. The evidence for this can be found in the fact that the calculated *ExpMinExp* and *ExpMinOpt* structures of **VIII** have *c*-lattice parameters that are 6.7 % and 7.4 % larger than the experimental value. By contrast, the calculated lattice parameters for the *ExpMinExp* and *ExpMinOpt* of **V** and **IYUPEX** display deviations of 3 % or less from the experimental values. Poor modelling of the C-F intermolecular contacts to CH - as indicated by a deviation of > 5 % in the C-F···H-Cangle of the cocrystal lattice energy minimum - is a likely contributor to the large deviation in the *c*-lattice parameter of the *ExpMinOpt* in **VIII**. This may also explain why the experimental structure of **VIII** was not predicted as the global minimum in lattice energy, given the importance of the C-F···H-C contacts in determining the relative stacking of the molecules in the cocrystal. Thus the same model for the lattice energy, which has proven reasonably accurate for **V** and **IYUPEX**, appears not to be as reliable for **VIII**.

Comparison of the cocrystal lattice energy with the sum of the lattice energies of its components has shown that **VIII** is approximately 9.57 kJ mol^{-1} more stable than the sum of the lattice energies of 4-cyanopyridine and 4-fluorobenzoic acid. Previous work by Issa *et al*[219] has shown that comparing the lattice energies of cocrystals with the sum of the lattice energies of their components can be suggestive of cocrystallisation but that the errors in the lattice energy estimates do not always cancel out. Thus, while we know from experimental evidence that **VIII** crystallises in preference to its components, we cannot be confident of the exact figure for the stabilisation energy of the cocrystal when compared to the stability of its component molecules.

5.5 Estimating the relative lattice energies of salts and cocrystals

One of the limitations of using rigid body lattice energy minimisations as applied to the prediction of salt and cocrystal structures is that such modelling is incapable of providing an accurate estimate of the energy penalty for proton transfer. Periodic electronic structure calculations are useful in this regard and the work presented in the last chapter (section 4.4.3) has shown that the energy penalty to relocate the acidic proton from a salt to a cocrystal correlates with the value of ΔpK_a as well as the presence/absence of any proton disorder in the crystal. It is important to appreciate however, that electronic structure calculations are computationally expensive and cannot be practically applied to many crystal structures without consuming a lot of time and effort.

For a given system, it would be more desirable to use the rigid body crystal energy landscapes of the salt and cocrystal molecular structures to provide a crude estimate of the relative lattice energies of the *Exp(Obs)MinOpt* and *Exp(Hyp)MinOpt*. As can be seen in the results for **V**, **VIII** and **IYUPEX**, the crystal energy landscape of the salt usually has a lattice energy scale that is approximately three times larger than that of the cocrystal. Thus if we are to compare the crystal energy landscapes of salts and cocrystals, we must offset the lattice energy of one solid form in order for it to be comparable to that of the other. One method of doing this is to compare the sum of the intramolecular energies of the acid-base pair with the sum of the intramolecular energies of the cation-anion pair. The intramolecular energies are taken from the single point quantum mechanical evaluation of the gas phase conformational energy minimum for each molecule or ion as calculated at the MP2/6-31G(d,p) level of theory. The single point MP2 calculations were performed at the optimised molecular conformations used in the search for hypothetical crystal structures. For a given system, the absolute difference in the sum of these intramolecular energies can be used to offset the lattice energy of the cocrystal and make it comparable to that of the salt or vice versa. From the outset, we note that the isolated molecule/ion energies on their own are not meaningful as the energy of the pair of molecules/ions in the crystal will be a function of the effects of neighbouring molecules/ions - something that cannot be sampled in isolated molecule/ion GAUSSIAN calculations. Furthermore, the energy barrier for

Solid form	Molecule/ion	E_{intra} / hartree	$\sum E_{intra}(i)$ / hartree	$\Delta \lvert \sum E_{intra}(i) \rvert$ / hartree	$\Delta \lvert \sum E_{intra}(i) \rvert$ / kJ mol^{-1}
V	Maleic acid	-454.50	-835.58	0.13	349.17
	4-dimethylaminopyridine	-381.07	-	-	-
	Maleate	-453.97	-835.45	-	-
	4-dimethylaminopyridinium	-381.48	-	-	-
VIII	4-fluorobenzoic acid	-518.62	-858.14	0.21	561.40
	4-cyanopyridine	-339.52	-	-	-
	4-fluorobenzoate	-518.05	-857.93	-	-
	4-cyanopyridinium	-339.88	-	-	-
IYUPEX	Isophthalic acid	-607.70	-855.23	0.19	501.64
	Pyridine	-247.52	-	-	-
	Isophthalate	-607.14	-855.03	-	-
	Pyridinium	-247.90	-	-	-

Table 5.5: Intramolecular energies of the isolated gas phase conformational minimum structures for the cation/anion or acid/base molecular structures used in crystal structure prediction for the salt and cocrystal structures of **V**, **VIII** and **IYUPEX**. $\sum E_{intra}(i)$ is the sum of the intramolecular energies of the acid-base or cation-anion pair. All quoted intramolecular energies are true at the MP2/6-31G(d,p) level of theory. Brown=Experimental molecular structures & Blue=Hypothetical molecular structures.

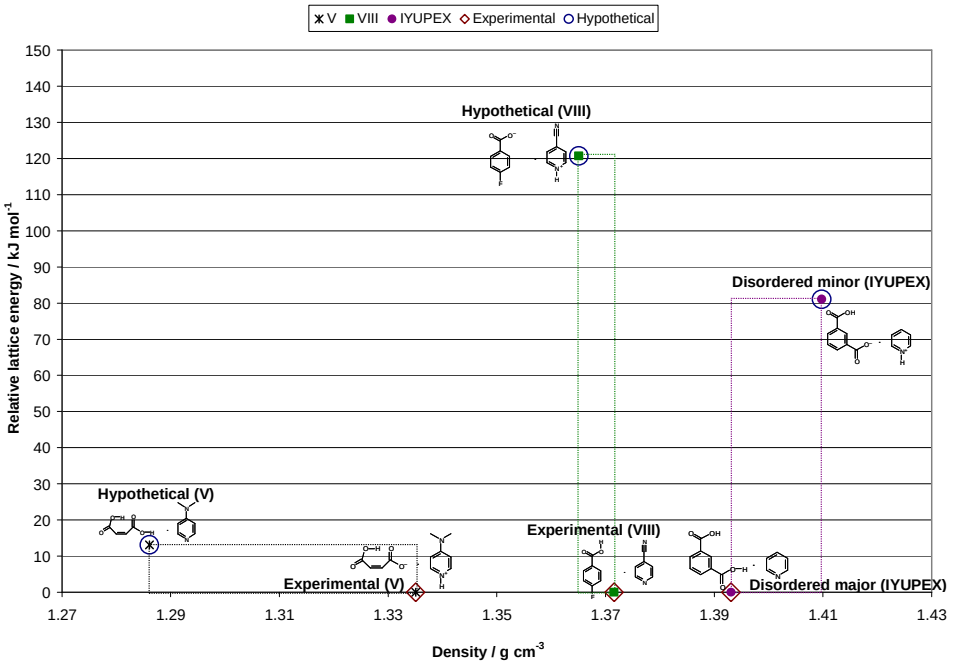


Figure 5.14: Plot of the relative lattice energy versus density for the experimental salt and hypothetical cocrystal or vice versa for systems **V**, **VIII** and **IYUPEX**. The relative lattice energies were estimated by using the intramolecular energies of the acid-base pair and cation-anion pair to offset the lattice energies of salts and cocrystals. Brown diamonds illustrate the experimental salt/cocrystal solid form while blue circles illustrate the hypothetical solid form generated following intermolecular proton transfer. For each system, the shaded box maps out the range of lattice energy and density between the salt and cocrystal.

proton transfer is a property of the crystal packing. Despite these caveats, it is worth validating how crude the relative energies obtained via this method are. The intramolecular energies of the cation-anion and acid-base pair of the salt and cocrystal molecular structures for **V**, **VIII** and **IYUPEX** are tabulated in Table 5.5.

For solid forms **V**, **VIII** and **IYUPEX**, if we offset the lattice energy of the cocrystal by an amount equal to the absolute difference in the sum of the intramolecular energies of the acid-base and cation-anion pairs, we can plot the relative lattice energy of the salt and cocrystal against the density for each structure and this is shown in Figure 5.14. The figure shows that for both **V** and **VIII**, which represent a pure salt and cocrystal system respectively, the observed solid form is energetically more stable than the hypothetical solid form generated following intermolecular proton transfer. For **V**, the experimental salt structure is 13.13 kJ mol^{-1} more stable than the hypothetical cocrystal whilst for **VIII,** the experimental cocrystal structure is approximately 120.73 kJ mol^{-1} more stable than the hypothetical salt. For the disordered **IYUPEX** structure, the cocrystal major product is estimated to be 81.06 kJ mol^{-1} more stable than the salt minor product. Thus for all systems, we estimate the observed structure (**V** and **VIII**) or disordered major product (cocrystal of **IYUPEX**) to be thermodynamically more stable than the hypothetical structure or disordered minor product. However we note that when we compare the relative lattice energies of the salt and cocrystal for the three systems, they are not in agreement with the value of ΔpK_a for the acid-base pair or indeed the effects of intermolecular proton disorder in the crystal. For example, we would expect **IYUPEX** to have the lowest relative lattice energy between the salt and cocrystal structures as a consequence of the intermolecular proton disorder between acid and base, but this is not reflected in the relative lattice energies (Figure 5.14). Instead the pure salt system, **V**, has the lowest relative lattice energy between the salt and cocrystal. It is also worth pointing out that for **IYUPEX**, the electronic structure calculations reported in the previous chapter have shown that the salt minor product is the thermodynamically more stable structure at 0 K, not the cocrystal major product as suggested (Figure 5.14) by the predicted lattice energies of the salt and cocrystal.

5.6 Conclusions

Crystallisation of an acid with a base can lead to a salt or cocrystal depending on the relative acidities of the two components, the solvent as well as the solubilities of the acid and base in the solvent used. In this chapter, we have carried out the computational crystal structure prediction of three systems: a salt (**V**) formed from complete proton transfer from acid to base, a disordered salt-cocrystal system (**IYUPEX**) and a cocrystal (**VIII**) that has the acidic proton fully localised on the acid. The experimental structure of **V** was predicted as the global minimum in the crystal energy landscape of the salt. This is an encouraging result which proves that current methods of computational crystal structure prediction can be successfully applied to the prediction of pyridinium carboxylate salts. The cocrystal major product of **IYUPEX** is ranked as the second most stable structure in the crystal energy landscape of the pyridine isophthalic acid cocrystal. Finally, the experimental structure of **VIII** is among the most stable predicted structures and is only 1.87 kJ mol^{-1} less stable than the predicted global minimum structure. Thus the results of crystal structure prediction using the experimental molecular structures of **V** and **VIII** as well as the molecular structure of the disordered major product for **IYUPEX**, suggest that all these structures are energetically favoured packing arrangements in their respective crystal energy landscapes.

By assuming the molecular structures of the salt and cocrystal in separate predictions, it was possible to assess the relative stability of each crystal structure in its respective crystal energy landscape. Computational crystal structure prediction of the hypothetical molecular structures of **V** and **VIII** have shown that the hypothetical salt/cocrystal structure is not as energetically favoured in the fictitious crystal energy landscape as the experimental salt/cocrystal structure is in the true crystal energy landscape. This was judged on the basis of the lattice energy of the hypothetical/experimental structure with respect to the predicted global minimum structure in the fictitious/true crystal energy landscapes. Similarly for **IYUPEX**, we found that the salt minor product was more energetically unfavoured in the salt energy landscape than was the case for the cocrystal major product in the crystal energy landscape of the cocrystal.

This work has shown that the results of computational crystal structure prediction can be sensitive to the atomic connectivity of the acidic proton in the input molecular

model. This was illustrated in the predictions for **V** and **IYUPEX**, which displayed significant re-ranking in the relative stabilities of structures based on the same hydrogen bond motifs when the crystal energy landscapes of the salt and cocrystal are compared. Furthermore, both systems displayed hydrogen bond motifs in the salt crystal energy landscape not seen in the cocrystal. The obvious caveat that goes with this is that the energy range of 10 kJ mol^{-1} considered in the analysis of the salt and cocrystal energy landscapes may have excluded these 'novel' motifs in one landscape but not in the other. Even if this is the case, it is further evidence of the effects of energetic re-ranking that is possible when an incorrect proton position is assumed in the predictions. For **VIII**, the molecular C_{2V} symmetry of both ions meant that there was only one version of the $R_2^2(7)$ and $D_1^1(2)$ motifs and this differed from the results for **V**, which were complicated by the intramolecular hydrogen bond of the maleate ion. Given that most organic molecules will have substituents that make it unlikely for the ions to have C_{2V} symmetry, future work on salts will be demanding of accurate models for the lattice energy as the interpretation of the crystal energy landscapes may be complicated by different versions of the $R_2^2(7)$ and $D_1^1(2)$ motifs. Certainly, the empirically derived force field used in this work has been successful for **V**, which is an example of a system that can display different versions of $R_2^2(7)$ and $D_1^1(2)$ motifs, but such success will not necessarily be replicated in other salt systems (*cf.* Chapter 7).

For solid form **V**, comparison of the packing discrepancy between the experimental structure and the lattice energy minima of the salt and cocrystal molecular structures has shown that the lattice energy minimum of the salt is a better match to the experimental structure. Similarly for **IYUPEX**, comparison of the packing discrepancies has shown that the cocrystal major product is a better match to the experimental structure when compared to the salt minor product. In the case of **VIII** we find that both the salt and cocrystal lattice energy minima are poor matches to the experimental structure. Thus it seems that when we have a reasonably accurate model for the forces between the molecules/ions, we can decide whether the correct molecular structure is a salt or a cocrystal on the basis of the packing discrepancy between the experimental structure and the calculated lattice energy minima of the salt and cocrystal molecular structures. It is important to note that such comparisons of the crystal packings could be performed even if the crystallography was unable to locate the

positions of the hydrogen atoms since the calculated root mean square deviations rely on the positions of non-hydrogen atoms alone.

For all systems, an estimate of the relative lattice energies of the calculated salt and cocrystal lattice energy minima was made by using the gas phase intramolecular energies of the acid-base and cation-anion pairs to offset the predicted lattice energies of the salt and cocrystal. In all cases, the experimental structure or disordered major product is estimated to be thermodynamically more stable than the hypothetical structure or disordered minor product. However when we compare across the three systems, the relative lattice energies of the salt and cocrystal are not consistent with the value of ΔpK_a for the acid-base pair or the effects of intermolecular proton disorder. Thus the crude estimates in the relative energies of salt-cocrystal pairs obtained via this method are unreliable for practical use and more accurate energies (albeit at greater computational cost) can be obtained via fixed cell periodic *ab initio* calculations.

5.7 Appendix

5.7.1 Predicted lattice energy minima of 4-dimethylaminopyridinium maleate

Structure ID	Elatt / kJ mol^{-1}	ρ / g cm^{-3}	Space Group	Hydrogen bond motifs	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
Exp(Obs)MinConOpt	-551.042	1.335	P21/c	R2, 2 (7)-Type 1	13.030	7.799	12.707	90	113.381	90
15	***-551.064***	***1.335***	***P21/c***	***R2, 2 (7)-Type 1***	***13.030***	***7.799***	***12.707***	***90***	***66.621***	***90***
27	-549.039	1.327	P21/c	R2, 2 (7) - Type 2	7.884	14.927	12.156	90	123.550	90
163	-547.362	1.300	P21/c	R2, 2 (7)-Type 1	12.799	7.705	12.367	90	86.697	90
22	-546.881	1.346	P21/c	R2, 2 (7) - Type 2	3.984	12.680	23.320	90	93.950	90
11	-546.288	1.321	P-1	D1, 1 (2)-Type 2	12.284	9.154	8.186	123.337	89.215	122.111
208	-545.902	1.286	P21/c	R2, 2 (7) - Type 2	7.167	12.102	14.349	90	81.442	90
45	-545.270	1.355	P21/c	R2, 2 (7) - Type 2	23.079	12.695	3.986	90	88.812	90
5	-545.230	1.298	P21/c	R2, 2 (7) - Type 2	9.777	17.750	7.720	90	65.492	90
59	-544.659	1.306	P-1	R2, 2 (7) - Type 2	9.733	9.611	7.872	73.224	113.512	114.926
2	-544.535	1.356	P21/c	D1, 1 (2)-Type 2	11.854	13.265	7.445	90	85.616	90
31	-544.277	1.352	P212121	R2, 2 (7) - Type 2	12.751	23.211	3.954	90	90	90
28	-543.838	1.313	P21/c	R2, 2 (7) - Type 2	10.120	8.024	15.123	90	78.95961	90
84	-543.774	1.316	P-1	R2, 2 (7) - Type 2	12.213	13.301	4.040	108.962	84.682	104.389
71	-543.506	1.324	P21/c	R2, 2 (7) - Type 2	3.927	22.879	13.581	90	78.35921	90
102	-543.422	1.319	P21/c	R2, 2 (7) - Type 2	12.890	3.955	24.857	90	71.17054	90
72	-543.419	1.334	P21/c	R2, 2 (7) - Type 2	3.926	22.455	13.684	90	79.66119	90
48	-543.387	1.347	P21/c	R2, 2 (7) - Type 2	4.002	20.329	14.552	90	82.841	90
101	-543.271	1.319	P21/c	R2, 2 (7) - Type 2	12.889	3.957	24.018	90	78.32313	90
193	-542.895	1.318	C2/c	R2, 2 (7)-Type 1	26.225	7.680	12.919	90	112.66754	90
83	-542.832	1.335	Pna21	R2, 2 (7) - Type 2	23.415	12.838	3.942	90	90	90
80	-542.641	1.344	Pna21	D1, 1 (2)-Type 1	24.853	11.913	3.977	90	90	90
860	-542.609	1.266	Pbcn	R2, 2 (7) - Type 2	12.837	24.423	7.973	90	90	90
8	-542.591	1.355	Pbca	D1, 1 (2)-Type 2	7.382	23.767	13.311	90	90	90

121	-542.537	1.315	P-1	R2, 2 (7) - Type 2	3.939	13.101	12.585	110.937	95.981	91.674
989	-542.532	1.299	P21/c	D1, 1 (2)-Type 2	5.220	12.040	21.204	90	113.966	90
124	-542.512	1.315	P-1	R2, 2 (7) - Type 2	12.789	3.939	13.101	91.675	68.862	101.853
126	-542.500	1.317	P-1	R2, 2 (7) - Type 2	6.334	8.127	12.504	89.257	100.935	107.862
183	-542.469	1.317	P21/c	R2, 2 (7) - Type 2	6.822	14.895	12.188	90	103.948	90
357	-542.379	1.292	P21/c	R2, 2 (7) - Type 2	8.217	6.339	24.708	90	107.894	90
91	-542.345	1.319	P-1	R2, 2 (7) - Type 2	3.916	13.794	12.225	112.193	89.406	100.664
165	-542.259	1.322	P212121	R2, 2 (7) - Type 2	6.862	14.829	11.762	90	90	90
839	-542.067	1.317	P-1	R2, 2 (7) - Type 2	12.341	13.739	3.921	78.019	91.929	112.288
240	-542.032	1.293	P-1	R2, 2 (7) - Type 2	6.325	13.760	8.667	100.570	63.200	114.634
18	-541.982	1.366	P21/c	D1, 1 (2)-Type 2	12.218	13.253	7.369	90	76.082	90
143	-541.896	1.326	P1	R2, 2 (7) - Type 2	7.314	6.381	6.934	96.755	83.112	70.355
589	-541.892	1.278	P-1	R2, 2 (7) - Type 2	8.286	6.374	12.905	109.566	98.782	75.108
962	-541.885	1.298	Pca21	R2, 2 (7) - Type 2	15.387	6.676	11.871	90	90	90
58	-541.862	1.360	P-1	D1, 1 (2)-Type 1	4.049	12.263	12.401	76.782	76.735	90.921
331	-541.711	1.280	P21/c	R2, 2 (7) - Type 2	12.273	7.979	12.628	90	90.012	90
99	-541.663	1.306	P-1	R2, 2 (7) - Type 2	6.880	10.469	11.000	107.481	74.257	126.143
118	-541.579	1.313	P21	R2, 2 (7) - Type 2	12.799	3.968	12.658	90	110.315	90
73	-541.570	1.344	Pna21	R2, 2 (7) - Type 2	23.430	4.011	12.531	90	90	90
90	-541.548	1.321	P-1	R2, 2 (7) - Type 2	3.917	13.816	12.159	112.108	89.878	80.080
20	-541.414	1.350	Pbcn	D1, 1 (2)-Type 2	13.451	7.372	23.634	90	90	90
342	-541.369	1.300	P21	R2, 2 (7) - Type 2	7.511	11.911	6.837	90	84.173	90
155	-541.337	1.322	P1	D1, 1 (2)-Type 2	7.570	10.044	4.321	112.885	82.143	90.077
385	-541.178	1.303	P-1	R2, 2 (7) - Type 2	7.289	12.566	8.673	51.050	80.235	86.960
129	-541.177	1.301	P21/c	R2, 2 (7) - Type 2	12.732	4.006	24.859	90	73.619	90
258	-541.174	1.306	P21/c	R2, 2 (7) - Type 2	6.374	8.201	23.597	90	79.183	90
505	-541.141	1.290	P21/c	R2, 2 (7) - Type 2	6.391	9.471	20.270	90	89.881	90

Table 5.6: The 50 most stable lattice energy minima in the crystal energy landscape of 4-dimethylaminopyridinium maleate. The predicted structure that corresponds to the *Exp(Obs)MinConOpt* of **V** is highlighted in bold.

5.7.2 Predicted lattice energy minima of 4-dimethylaminopyridine maleic acid

Structure ID	E_{latt} / kJ mol^{-1}	ρ / g cm^{-3}	Space Group	Hydrogen bond motif	*a* (Å)	*b* (Å)	*c* (Å)	α (°)	β (°)	γ (°)
Exp(Hyp)MinConOpt	-188.740	1.286	P21/c	R2, 2 (7)-Type 1	13.193	7.934	12.976	90	115.062	90
30	-191.144	1.280	P21/c	D1, 1 (2)-Type 1	11.707	7.096	16.793	90	117.625	90
34	-190.736	1.278	P21/c	D1, 1 (2)-Type 1	11.752	7.083	16.821	90	117.809	90
12	-190.648	1.315	P-1	R2, 2 (7)-Type 1	8.118	8.781	8.916	86.293	98.101	106.959
57	-190.617	1.277	P21/c	R2, 2 (7)-Type 1	11.753	7.086	16.850	90	117.940	90
24	-190.467	1.312	P-1	D1, 1 (2)-Type 1	8.065	9.004	8.758	86.645	73.210	82.093
19	-190.003	1.328	P21/c	R2, 2 (7)-Type 1	6.388	27.926	7.481	90	116.758	90
23	-189.713	1.309	P-1	D1, 1 (2)-Type 1	8.951	8.158	8.742	107.262	86.995	97.570
53	-189.662	1.307	P-1	R2, 2 (7)-Type 1	8.111	8.693	9.046	87.674	82.617	73.223
256	-189.510	1.279	P21/c	D1, 1 (2)-Type 1	19.961	7.135	8.874	90	78.130	90
13	-189.273	1.305	Cc	D1, 1 (2)-Type 1	9.463	10.915	13.617	90	59.557	90
329	-189.198	1.277	P21/c	R2, 2 (7)-Type 1	19.996	7.114	8.893	90	78.454	90
322	-189.113	1.279	P21/c	D1, 1 (2)-Type 1	8.866	7.134	20.158	90	103.997	90
422	-189.014	1.285	P21/c	R2, 2 (7)-Type 1	12.979	7.908	13.221	90	65.130	90
36	-188.792	1.287	P21/c	R2, 2 (7)-Type 1	7.819	15.295	11.922	90	59.618	90
419	-188.768	1.276	P21/c	R2, 2 (7)-Type 1	8.896	7.110	20.243	90	104.413	90
585	-188.752	1.305	P21/c	R2, 2 (7)-Type 1	10.219	11.411	10.445	90	84.608	90
41	-188.745	1.320	P21/c	R2, 2 (7)-Type 1	7.438	28.015	6.366	90	64.660	90
380	***-188.736***	***1.286***	***P21/c***	D1, 1 (2)-Type 1	***13.193***	***7.934***	***12.976***	***90***	***64.939***	***90***
16	-188.697	1.304	Cc	D1, 1 (2)-Type 1	9.458	10.929	13.562	90	59.954	90
651	-188.578	1.304	P21/c	D1, 1 (2)-Type 1	10.206	11.451	10.423	90	85.123	90
603	-188.521	1.304	P21/c	R2, 2 (7)-Type 1	10.211	11.576	10.301	90	85.100	90
40	-188.488	1.322	P21/c	D1, 1 (2)-Type 1	7.344	28.013	6.384	90	65.666	90
54	-188.215	1.286	Pbca	D1, 1 (2)-Type 1	28.843	10.511	8.121	90	90	90
51	-188.208	1.285	P21/c	R2, 2 (7)-Type 1	7.830	15.291	11.926	90	59.557	90
360	-188.172	1.312	P-1	R2, 2 (7)-Type 1	7.201	12.244	8.690	55.210	99.949	84.939
379	-188.090	1.310	P-1	D1, 1 (2)-Type 1	7.213	8.699	10.215	80.702	74.290	80.718

373	-188.017	1.312	P-1	R2, 2 (7)-Type 1	8.694	7.201	10.216	74.609	99.522	99.694
32	-188.006	1.314	P21	D1, 1 (2)-Type 1	6.303	7.593	12.783	90	100.260	90
979	-187.997	1.277	P21/c	R2, 2 (7)-Type 1	7.159	11.769	18.598	90	127.756	90
72	-187.975	1.282	P21/c	R2, 2 (7)-Type 1	10.701	6.754	17.308	90	80.623	90
346	-187.950	1.311	P-1	D1, 1 (2)-Type 1	10.252	8.701	7.201	98.989	73.435	99.193
121	-187.924	1.284	P-1	R2, 2 (7)-Type 1	11.184	6.359	9.877	111.759	108.387	78.392
11	-187.891	1.311	P-1	D1, 1 (2)-Type 1	6.507	12.924	7.982	93.809	70.681	107.548
244	-187.890	1.304	P-1	D1, 1 (2)-Type 1	11.741	7.749	7.373	66.743	91.582	81.512
95	-187.858	1.282	P21/c	D1, 1 (2)-Type 1	7.740	15.317	12.036	90	59.899	90
233	-187.812	1.283	P21/c	R2, 2 (7)-Type 1	6.972	15.269	11.876	90	102.666	90
5	-187.779	1.303	Pbca	D1, 1 (2)-Type 1	7.838	28.049	11.045	90	90	90
705	-187.760	1.277	P21/c	D1, 1 (2)-Type 1	7.146	11.777	18.507	90	52.719	90
210	-187.674	1.276	P21/c	R2, 2 (7)-Type 1	12.080	8.120	12.734	90	83.283	90
38	-187.576	1.314	P21	D1, 1 (2)-Type 1	15.223	7.587	6.323	90	124.472	90
76	-187.563	1.284	Pbca	D1, 1 (2)-Type 1	8.131	28.753	10.545	90	90	90
6	-187.530	1.303	Pbca	R2, 2 (7)-Type 1	28.017	10.938	7.923	90	90	90
122	-187.510	1.280	P-1	R2, 2 (7)-Type 1	11.905	6.341	9.735	68.653	84.373	110.676
97	-187.494	1.283	Pbca	R2, 2 (7)-Type 1	28.792	10.440	8.209	90	90	90
10	-187.395	1.303	Pbca	D1, 1 (2)-Type 1	27.994	10.924	7.944	90	90	90
65	-187.386	1.313	P21	D1, 1 (2)-Type 1	12.767	7.594	6.320	90	79.523	90
140	-187.379	1.315	P21/c	D1, 1 (2)-Type 1	6.362	28.083	7.440	90	64.847	90
218	-187.352	1.280	P21/c	R2, 2 (7)-Type 1	12.073	8.049	12.791	90	83.994	90
118	-187.341	1.280	P21/c	D1, 1 (2)-Type 1	7.757	15.310	12.040	90	59.813	90
688	-187.326	1.276	P21/c	R2, 2 (7)-Type 1	18.679	11.801	7.186	90	51.515	90

Table 5.7: The 50 most stable lattice energy minima in the crystal energy landscape of the hypothetical 4-dimethylaminopyridine maleic acid cocrystal. The predicted structure that corresponds to the *Exp(Hyp)MinConOpt* is highlighted in bold.

5.7.3 Predicted lattice energy minima of pyridine isophthalic acid

Structure ID	E_{latt} / kJ mol^{-1}	ρ / g cm^{-3}	Space Group	Hydrogen bond motif	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
Exp-cocrystal	-192.157	1.393	P21/c	C1,1(8)-COLUMN 1	3.762	17.721	17.554	90	87.450	90
173	-192.324	1.359	P-1	R2,2(8)-DIMER	14.410	7.167	6.105	74.138	87.222	97.337
2	***-192.165***	***1.393***	***P21/c***	**C1,1(8)-COLUMN 1**	***3.762***	***17.721***	***18.789***	***90***	***111.036***	***90***
456	-191.974	1.380	P-1	R2,2(8)-DIMER	27.624	7.002	3.817	68.434	82.780	114.496
39	-191.968	1.380	P-1	R2,2(8)-DIMER	23.772	3.817	7.002	68.434	90.617	87.627
643	-191.577	1.385	P21/c	R2,2(8)-DIMER	6.107	3.766	51.494	90	83.379	90
652	-191.571	1.385	P21/c	R2,2(8)-DIMER	6.107	3.766	51.491	90	96.627	90
144	-191.548	1.377	P21/c	R2,2(8)-DIMER	26.096	3.828	12.734	90	68.371	90
528	-190.844	1.376	P21/c	R2,2(8)-DIMER	3.790	6.092	55.440	90	112.369	90
540	-190.826	1.376	P21/c	R2,2(8)-DIMER	3.790	6.092	51.613	90	83.401	90
871	-190.597	1.336	P21/c	R2,2(8)-DIMER	4.972	7.226	33.988	90	93.568	90
887	-190.593	1.336	P21/c	R2,2(8)-DIMER	4.972	7.226	53.905	90	141.002	90
7	-190.565	1.392	P212121	C1,1(8)-COLUMN 1	17.944	3.793	17.190	90	90	90
489	-190.321	1.340	P-1	R2,2(8)-DIMER	7.206	17.391	5.000	88.397	96.048	77.524
156	-190.320	1.336	P-1	R2,2(8)-DIMER	16.434	7.015	7.147	99.985	105.457	123.242
145	-190.319	1.336	P-1	R2,2(8)-DIMER	7.015	7.147	12.806	83.043	75.374	80.015
227	-189.833	1.366	C2/c	R2,2(8)-DIMER	13.138	3.767	48.216	90	91.926	90
228	-189.825	1.366	C2/c	R2,2(8)-DIMER	13.138	3.767	49.546	90	103.445	90
757	-189.808	1.370	P21/c	R2,2(8)-DIMER	26.310	3.820	12.697	90	111.337	90
133	-189.773	1.354	C2/c	R2,2(8)-DIMER	49.898	3.873	13.225	90	109.673	90
419	-189.753	1.354	P-1	R2,2(8)-DIMER	3.737	6.163	31.123	62.701	106.938	89.821
202	-189.597	1.364	P21/c	R2,2(8)-DIMER	3.823	48.274	6.570	90	80.085	90
211	-189.578	1.364	P21/c	R2,2(8)-DIMER	3.823	48.272	7.009	90	112.581	90
575	-189.485	1.325	P21/c	R2,2(8)-DIMER	14.076	7.197	12.196	90	95.576	90
9	-189.426	1.396	Pna21	C1,1(8)-COLUMN 1	18.378	17.223	3.685	90	90	90
585	-189.421	1.337	P21/c	R2,2(8)-DIMER	4.945	34.610	8.163	90	60.649	90
981	-189.380	1.347	P-1	R2,2(8)-DIMER	11.532	8.225	7.126	92.884	105.547	68.446
371	-189.321	1.333	P21/c	R2,2(8)-DIMER	11.567	15.956	6.717	90	80.268	90

447	-189.221	1.361	C2/c	R2,2(8)-DIMER	50.249	3.724	12.852	90	84.582	90
439	-189.220	1.361	C2/c	R2,2(8)-DIMER	54.236	3.724	12.852	90	67.267	90
909	-189.177	1.318	C2/c	R2,2(8)-DIMER	29.789	7.205	12.213	90	70.564	90
991	-189.099	1.354	P21/c	R2,2(8)-DIMER	6.257	51.692	3.721	90	87.781	90
984	-189.090	1.354	P21/c	R2,2(8)-DIMER	3.721	51.689	6.257	90	87.776	90
13	-189.002	1.361	P-1	C1,1(8)-COLUMN 2	8.907	7.271	10.551	69.623	86.945	108.238
417	-188.748	1.334	P21/c	R2,2(8)-DIMER	8.818	7.203	19.249	90	87.359	90

Table 5.8: The 34 most stable lattice energy minima in the crystal energy landscape of pyridine isophthalic acid (1:1). The predicted structure matching the cocrystal major product of **IYUPEX** is highlighted in bold.

5.7.4 Predicted lattice energy minima of pyridinium isophthalate

Structure ID	Elatt / kJ mol^{-1}	ρ / g cm^{-3}	Space Group	Hydrogen bond motif	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
Exp-salt	-612.734	1.410	P21/c	C1,1(8)-COLUMN 1	3.843	17.579	17.130	90	86.795	90
307	-619.584	1.400	P21/c	C1,1(8)-RIBBON 2	12.226	8.353	11.521	90	81.467	90
1	-619.204	1.400	P21/c	C1,1(8)-RIBBON 2	11.520	8.354	12.226	90	81.468	90
4	-617.664	1.426	P21/c	C1,1(8)-RIBBON 1	11.755	8.024	12.248	90	81.401	90
6	-616.596	1.403	P21/c	C1,1(8)-RIBBON 1	12.045	7.868	14.740	90	56.230	90
2	-616.257	1.385	Pna21	C1,1(8)-RIBBON 1	22.135	10.019	5.303	90	90	90
25	-615.285	1.407	P21/c	C1,1(8)-RIBBON 3	5.020	17.916	15.823	90	54.461	90
10	-614.884	1.412	P212121	C1,1(8)-COLUMN 1	16.921	16.697	4.082	90	90	90
11	-613.738	1.419	P-1	R2,2(8)-DIMER	10.789	3.818	14.111	87.492	93.148	82.055
19	-613.668	1.418	Pna21	C1,1(8)-COLUMN 1	17.671	16.821	3.863	90	90	90
371	-612.882	1.335	P21/c	R2,2(8)-DIMER	8.055	5.912	26.411	90	104.011	90
5	***-612.739***	***1.410***	***P21/c***	***C1,1(8)-COLUMN 1***	***3.843***	***17.579***	***17.130***	***90***	***86.794***	***90***
90	-612.548	1.387	C2/c	R2,2(8)-DIMER	21.425	3.800	29.528	90	102.291	90
237	-612.034	1.380	P21/c	R2,2(8)-DIMER	17.138	3.789	20.290	90	116.363	90
50	-611.635	1.427	P21/c	C1,1(8)-RIBBON 3	3.962	16.252	17.812	90	95.400	90
12	-611.288	1.391	P21/c	R2,2(8)-DIMER	9.975	4.087	28.727	90	88.947	90
332	-611.280	1.357	P21/c	R2,2(8)-DIMER	11.013	4.825	22.739	90	83.447	90
91	-611.223	1.383	P21/c	R2,2(8)-DIMER	11.026	3.999	27.579	90	104.432	90
3	-610.927	1.374	P212121	C1,1(8)-RIBBON 1	22.685	5.511	9.482	90	90	90
43	-610.783	1.383	P21/c	R2,2(8)-DIMER	3.924	27.287	12.615	90	119.308	90
46	-610.610	1.373	P21/c	C1,1(8)-RIBBON 3	5.126	13.021	19.168	90	112.039	90
393	-610.482	1.352	C2/c	R2,2(8)-DIMER	28.302	11.884	7.182	90	94.083	90
27	-610.379	1.402	P21/c	R2,2(8)-DIMER	14.346	3.876	21.365	90	102.023	90
524	-610.354	1.397	C2/c	R2,2(8)-DIMER	20.427	3.730	31.213	90	101.236	90
162	-610.271	1.401	C2/c	C1,1(8)-RIBBON 2	20.015	7.560	15.572	90	80.647	90
22	-610.209	1.391	P-1	R2,2(8)-DIMER	3.751	15.875	11.173	105.889	68.197	104.619
20	-610.146	1.354	P21/c	C1,1(8)-RIBBON 3	12.144	7.791	12.726	90	87.653	90
192	-609.826	1.354	P-1	R2,2(8)-DIMER	11.287	8.080	6.881	88.126	78.640	77.918

165	-609.675	1.383	P21/c	R2,2(8)-DIMER	20.419	3.868	16.186	90	67.158	90
59	-609.542	1.412	P21/c	C1,1(8)-RIBBON 3	3.818	17.817	17.008	90	94.518	90
292	-609.511	1.305	P21/c	C1,1(8)-RIBBON 3	10.018	14.354	10.110	90	120.869	90
48	-609.355	1.366	P-1	R2,2(8)-DIMER	10.794	8.501	7.283	84.071	74.334	67.930
168	-609.267	1.393	P21/c	C1,1(4)-CATEMER	15.271	3.856	20.399	90	103.301	90
373	-609.248	1.325	P21/c	C1,1(8)-RIBBON 3	15.164	15.738	5.469	90	70.422	90
93	-609.103	1.381	P21/c	R2,2(8)-DIMER	16.446	3.971	20.230	90	116.757	90

Table 5.9: The 34 most stable lattice energy minima in the crystal energy landscape of pyridinium isophthalate (1:1). The predicted structure matching the salt minor product of **IYUPEX** is highlighted in bold.

5.7.5 Predicted lattice energy minima of 4-cyanopyridine 4-fluorobenzoic acid

Structure ID	E_{latt} / kJ mol^{-1}	ρ / g cm^{-3}	Space Group	Hydrogen bond motif	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
Exp(Obs)MinOpt	-171.968	1.372	C2/c	R2, 2(7)-DIMER	12.661	7.369	25.448	90	85.076	90
1	-173.835	1.374	P-1	R2, 2(7)-DIMER	13.585	7.032	7.158	71.807	65.605	85.966
390	-173.304	1.367	P-1	D1, 1(2)-DISCRETE	27.423	3.741	6.945	57.888	81.930	91.502
296	-173.183	1.368	P21/c	D1, 1(2)-DISCRETE	3.745	5.870	54.306	90	83.477	90
179	-173.178	1.368	P21/c	D1, 1(2)-DISCRETE	5.877	3.746	54.812	90	79.405	90
34	-173.097	1.368	P21/c	D1, 1(2)-DISCRETE	3.747	53.914	6.944	90	57.733	90
87	-172.900	1.364	P21/c	D1, 1(2)-DISCRETE	3.755	5.859	55.762	90	75.763	90
2	-172.836	1.373	P-1	R2, 2(7)-DIMER	7.140	7.061	12.492	87.728	97.065	71.595
240	-172.808	1.364	P21	D1, 1(2)-DISCRETE	3.757	5.851	27.228	90	96.640	90
137	-172.733	1.365	P212121	D1, 1(2)-DISCRETE	54.025	3.757	5.857	90	90	90
360	-172.624	1.365	Pna21	D1, 1(2)-DISCRETE	54.053	3.759	5.850	90	90	90
3	-172.403	1.354	P21/c	R2, 2(7)-DIMER	12.193	4.085	24.361	90	80.760	90
5	-172.347	1.365	P21/c	R2, 2(7)-DIMER	12.380	7.033	13.693	90	85.668	90
7	-172.330	1.376	P-1	R2, 2(7)-DIMER	3.646	12.302	14.358	98.344	111.373	81.083
321	-172.110	1.359	P21/c	D1, 1(2)-DISCRETE	5.930	3.739	55.012	90	78.081	90
17	-172.108	1.382	P21/c	R2, 2(7)-DIMER	13.505	3.648	23.932	90	84.737	90
62	-171.985	1.373	P21/c	R2, 2(7)-DIMER	18.931	3.629	17.259	90	84.893	90
37	***-171.969***	***1.372***	***C2/c***	***R2, 2(7)-DIMER***	***12.661***	***7.369***	***25.447***	***90***	***85.075***	***90***
311	-171.952	1.360	P21	D1, 1(2)-DISCRETE	5.930	3.741	27.489	90	101.996	90
687	-171.938	1.359	P212121	D1, 1(2)-DISCRETE	53.913	5.917	3.741	90	90	90
932	-171.834	1.360	Pna21	D1, 1(2)-DISCRETE	53.832	5.920	3.744	90	90	90
8	-171.784	1.372	C2/c	D1, 1(2)-DISCRETE	26.253	3.650	24.732	90	86.036	90
569	-171.775	1.376	C2/c	D1, 1(2)-DISCRETE	25.770	3.663	25.117	90	84.053	90
187	-171.610	1.373	P21/c	D1, 1(2)-DISCRETE	12.603	6.099	15.469	90	83.428	90
717	-171.529	1.376	P21/c	R2, 2(7)-DIMER	26.572	3.617	12.311	90	85.147	90
10	-171.501	1.358	C2/c	R2, 2(7)-DIMER	24.742	7.039	13.751	90	85.822	90
48	-171.477	1.373	P21/c	R2, 2(7)-DIMER	13.359	3.615	24.564	90	84.877	90
83	-171.200	1.357	C2/c	D1, 1(2)-DISCRETE	54.279	3.714	12.025	90	99.613	90

122	-171.182	1.355	C2/c	D1, 1(2)-DISCRETE	53.426	3.717	12.069	90	92.779	90
533	-171.128	1.385	P-1	R2, 2(7)-DIMER	9.820	6.613	10.440	67.660	80.640	103.489
11	-171.089	1.383	P21/c	R2, 2(7)-DIMER	17.313	3.590	18.934	90	85.075	90
66	-171.086	1.383	P21	R2, 2(7)-DIMER	13.365	3.590	12.267	90	95.140	90
85	-171.033	1.331	P21/c	D1, 1(2)-DISCRETE	3.691	44.306	8.942	90	56.464	90
503	-170.802	1.348	C2	D1, 1(2)-DISCRETE	54.593	3.746	5.884	90	91.651	90
767	-170.797	1.348	P21212	D1, 1(2)-DISCRETE	54.659	5.877	3.746	90	90	90
637	-170.758	1.382	P21/c	D1, 1(2)-DISCRETE	12.695	4.002	23.247	90	83.795	90
15	-170.731	1.384	P-1	R2, 2(7)-DIMER	17.049	6.616	6.800	123.302	103.619	98.447
6	-170.722	1.367	C2/c	D1, 1(2)-DISCRETE	52.451	3.708	12.470	90	78.125	90
176	-170.638	1.363	P21/c	R2, 2(7)-DIMER	12.544	3.706	26.077	90	101.043	90
63	-170.574	1.364	P21/c	D1, 1(2)-DISCRETE	17.587	3.730	18.293	90	82.371	90
450	-170.477	1.367	P21/c	D1, 1(2)-DISCRETE	7.392	6.088	26.717	90	99.218	90
30	-170.382	1.363	P21/c	R2, 2(7)-DIMER	25.575	3.716	12.544	90	93.190	90

Table 5.10: The 41 most stable lattice energy minima in the crystal energy landscape of 4-cyanopyridine 4-fluorobenzoic acid (1:1). The predicted structure that matches the *Exp(Obs)MinOpt* of **VIII** is highlighted in bold.

5.7.6 Predicted lattice energy minima of 4-cyanopyridinium 4-fluorobenzoate

Structure ID	E_{latt} / kJ mol^{-1}	ρ / g cm^{-3}	Space Group	Hydrogen bond motif	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
Exp(Hyp)MinOpt	-612.637	1.365	C2/c	R2, 2(7)-DIMER	12.677	7.629	24.577	90	90.852	90
1	-621.096	1.411	P21/c	R2, 2(7)-DIMER	7.614	6.702	22.548	90	91.975	90
2	-619.509	1.399	P21	R2, 2(7)-DIMER	11.428	6.649	7.677	90	96.288	90
551	-618.965	1.400	P21/c	R2, 2(7)-DIMER	7.121	22.238	7.341	90	94.500	90
7	-618.821	1.387	Pna21	R2, 2(7)-DIMER	22.002	7.739	6.868	90	90	90
377	-618.699	1.408	P-1	D1, 1(2)-DISCRETE	6.981	7.338	13.592	56.070	86.324	86.363
3	-618.451	1.402	Pna21	R2, 2(7)-DIMER	15.421	11.598	6.468	90	90	90
567	-618.396	1.410	P21/c	D1, 1(2)-DISCRETE	7.433	11.144	19.646	90	135.004	90
18	-618.145	1.398	P21/c	R2, 2(7)-DIMER	6.782	7.568	22.617	90	91.747	90
6	-617.878	1.395	P21/c	R2, 2(7)-DIMER	7.718	6.823	22.198	90	95.827	90
20	-617.622	1.380	P212121	R2, 2(7)-DIMER	7.789	6.737	22.402	90	90	90
144	-617.281	1.370	P21/c	D1, 1(2)-DISCRETE	7.391	11.266	14.616	90	76.618	90
156	-616.291	1.382	P-1	R2, 2(7)-DIMER	7.720	12.066	7.731	81.259	55.533	85.087
9	-616.289	1.382	P-1	R2, 2(7)-DIMER	7.198	7.731	12.066	81.259	85.905	62.156
4	-616.277	1.370	P21/c	D1, 1(2)-DISCRETE	7.392	11.265	14.618	90	103.321	90
32	-615.412	1.391	C2/c	D1, 1(2)-DISCRETE	10.695	9.597	22.834	90	95.468	90
68	-615.238	1.391	P-1	R2, 2(7)-DIMER	12.452	7.219	7.357	115.592	98.829	78.685
14	-615.038	1.380	C2/c	R2, 2(7)-DIMER	13.583	7.204	24.349	90	99.321	90
13	-614.864	1.383	P212121	R2, 2(7)-DIMER	7.682	6.610	23.094	90	90	90
63	-614.440	1.385	P21/c	R2, 2(7)-DIMER	12.394	6.954	13.627	90	85.799	90
84	-614.294	1.384	C2/c	D1, 1(2)-DISCRETE	24.531	7.029	13.622	90	86.337	90
38	-613.767	1.376	P-1	R2, 2(7)-DIMER	7.366	11.735	6.978	84.345	93.977	100.457
28	-613.721	1.373	P21	R2, 2(7)-DIMER	7.744	6.580	11.589	90	89.920	90
27	-613.699	1.377	Pna21	R2, 2(7)-DIMER	23.360	7.652	6.591	90	90	90
80	-613.486	1.386	P-1	R2, 2(7)-DIMER	7.439	12.361	7.036	90.067	65.318	85.212
59	-613.117	1.368	C2/c	R2, 2(7)-DIMER	12.624	7.415	25.416	90	85.433	90
30	-613.035	1.399	P21	D1, 1(2)-DISCRETE	8.060	6.035	12.377	90	105.612	90
22	***-612.782***	***1.364***	***C2/c***	***R2, 2(7)-DIMER***	***12.671***	***7.622***	***24.638***	***90***	***90.855***	***90***

60	-612.726	1.376	P-1	R2, 2(7)-DIMER	7.615	7.163	12.361	83.725	90.211	61.787
67	-612.672	1.336	P21/c	D1, 1(2)-DISCRETE	14.258	6.953	13.397	90	113.941	90
90	-612.334	1.359	P-1	R2, 2(7)-DIMER	7.058	7.324	12.983	98.209	90.377	64.155
373	-612.234	1.367	C2/c	R2, 2(7)-DIMER	12.687	7.633	24.518	90	88.950	90
87	-612.231	1.366	C2/c	R2, 2(7)-DIMER	12.685	7.632	27.814	90	118.140	90
385	-612.222	1.345	C2/c	R2, 2(7)-DIMER	13.060	7.226	28.392	90	115.788	90
91	-612.220	1.345	C2/c	R2, 2(7)-DIMER	13.060	7.226	25.574	90	88.413	90
25	-612.066	1.373	P21/c	R2, 2(7)-DIMER	7.649	6.652	23.597	90	79.819	90
71	-611.649	1.377	P21/c	R2, 2(7)-DIMER	7.009	11.107	15.132	90	89.497	90
70	-611.608	1.356	P21/c	R2, 2(7)-DIMER	7.494	12.513	13.736	90	68.243	90
53	-611.324	1.356	P21/c	R2, 2(7)-DIMER	7.271	22.603	7.345	90	82.306	90
112	-611.249	1.380	P21/c	D1, 1(2)-DISCRETE	11.345	8.378	12.691	90	77.019	90
320	-611.102	1.352	P21/c	R2, 2(7)-DIMER	7.162	26.997	7.051	90	118.364	90
170	-611.071	1.379	P-1	R2, 2(7)-DIMER	13.664	7.110	6.976	62.492	99.575	88.055

Table 5.11: The 41 most stable lattice energy minima in the crystal energy landscape of the 4-cyanopyridinium 4-fluorobenzoate salt. The predicted structure that matches the *Exp(Hyp)MinOpt* is highlighted in bold.

6 Crystal structure prediction of amantadine hydrochloride and memantine hydrochloride

6.1 Introduction

Amantadine hydrochloride, **IX**, and memantine hydrochloride, **X**, are salt systems (Scheme 6.1) that display pharmacological activity. Amantadine hydrochloride is an antiviral drug that is primarily used for the treatment of flu caused by the influenza A virus[220]. The drug has also been linked to the treatment of Parkinson's disease through its action of stimulating the release of dopamine in the brain. The increased levels of dopamine are effective in reducing the tremors experienced by people suffering from Parkinson's disease. Its mode of action against the influenza A virus is less clear but is thought to involve the inhibition of virus replication. The drug is marketed under several trade names (Symmetrel®, Lysovir®) and is listed in the National Health Service (NHS) catalogue of approved medicines. Memantine hydrochloride is an approved drug for the treatment of moderate to severe forms of Alzheimer's disease. The disease is known to account for nearly half of all known dementia cases and like all neurodegenerative disorders, the risk of developing Alzheimer's disease increases with age. The current estimate of worldwide sufferers of the disease stands at 15 million, a number expected to treble by the year 2050[221]. Memantine hydrochloride (Nemanda®, Axura®, Ebixa®) was approved[222] for the treatment of Alzheimer's disease by the European Medicines Agency in 2002 and by the U.S. Food and Drug Administration in 2003. The use of **X** in the treatment of Alzheimer's disease stems from the ability of the free base to act as an antagonist to the *N*-methyl-*D*-aspartate (NMDA) receptor

IX **X**

Scheme 6.1: Sketch of the molecular structures for amantadine hydrochloride (1-aminoadamantane hydrochloride), **IX**, and memantine hydrochloride (3, 5-dimethyl-1-aminoadamantane hydrochloride, **X**). The known crystal structure of **IX** has the CSD refcode **FINVAZ** while the internal structure code for **X** is *tinr3c0m*.

found on neurons[223]. The NMDA receptor is a type of ligand-gated ion channel whose opening and closing is regulated by the levels of glutamate in extracellular fluid. The receptor has been identified as having an important physiological role in the processes of learning and memory retention. There is a growing body of evidence[223,224] suggesting a link between over-activation of the NMDA receptor due to excessive levels of glutamate and the development of neurodegenerative diseases such as Alzheimer's. In the UK, the National Institute for Health and Clinical Excellence (NICE) has not yet approved **X** for routine use within the NHS. The official stance[225] of NICE is that **X** 'could not reasonably be considered a cost-effective therapy' for patients suffering from moderate to severe Alzheimer's disease as judged on the basis of the available evidence from clinical trials.

Given the wide use of the chloride anion in salts of pharmaceutically acceptable molecules[172], it is important to validate the ability of computational models to predict the crystal structures of chloride salts. This chapter describes the results of a joint experimental and computational search for the polymorphs of **IX** and **X**. All the experimental work presented was performed by the group of Prof. Alistair Florence (Strathclyde University). Amantadine hydrochloride, **IX**, has a known[226] crystal structure (**FINVAZ**) in the Cambridge Structural Database (CSD) and previous work has suggested[227] a phase transition at 124 K. We report in this chapter the structure of **IX** at a temperature of 105 K (internal structure code: *aman_103_0m*). For anhydrous memantine hydrochloride, **X,** experiments at Strathclyde have led to one fully characterised crystal structure (internal structure code: *tinr3c0m*) and all references to the "experimental structure" refer to *tinr3c0m***.** A monohydrate crystal structure of **X** has also been discovered in the experimental screens. A search of the literature revealed a patent[228] claiming the existence of two "polymorphic" forms of **X** based on the combined results of powder X-ray diffraction, infrared spectroscopy and differential scanning calorimetry experiments. Careful inspection of the patent revealed Form II to be a hydrate rather than a strict polymorph of **X**. Comparison of the X-ray powder diffraction patterns for the patent Form I structure and that simulated (λ=1.54056 Å) from the crystal structure of *tinr3c0m* showed good agreement. This comparison is not included in the results section of this chapter because the patent documentation is unclear of the type of radiation used to collect the X-ray powder diffraction data, thereby precluding a meaningful comparison with the crystal structure of *tinr3c0m*.

During the course of our work, a crystal structure of memantine hydrochloride with 10 % water occupancy had been published[229]. Comparison with the *tinr3c0m* structure shows that the structures are the same apart from the inclusion of water in the refinement of the newly published structure.

6.2 Method

6.2.1 Computational crystal structure prediction

The experimental conformations of the amantadinium and memantinium cations were extracted from the known anhydrous crystal structures of **IX** and **X** and used in separate gas phase geometry optimisations at the MP2/6-31G (d,p) level of theory using the program GAUSSIAN03[119]. Using the *ab initio* optimised molecular conformations of the amantadinium and memantinium cations, CrystalPredictor[88] was used to generate hypothetical crystal structures of **IX** and **X**. Searches were performed in the space groups $P1$, $P\bar{1}$, $P2_1$, $P2_1/c$, $P2_12_12$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$, $Pbca$, $Pbcn$, Cc, $C2$, $C2/c$, $P2_1/m$, $P3_121$, $P3_221$, $R3c$, $R3$, $P\bar{3}$, $R\bar{3}$, $P6_1$, $P6_3$ and $P6_3/m$. A total of 80,000 trial crystal structures were generated for **IX** while the search for hypothetical crystal structures of **X** generated 101,500 trial crystal structures. Convergence was assessed on the basis of the number of times each of the low energy structures was found in the search. For both **IX** and **X**, the global minimum structure was found more than fifteen times by CrystalPredictor and this was sufficient to confirm that the search had reached completion. An *exp*-6 atom-atom potential was used to calculate the dispersion-repulsion contributions towards the lattice energy. The potential parameters for the atoms C, H_C, N and H_N were taken from the W99[107] force field while the parameters for the Cl^- ion were taken from the fitting of Hejczyk[230]. Validation of this choice of potential parameters can be found in the appendix (see section 6.5.4). Charges on the individual atoms of a cation were derived by fitting[231] to the electrostatic potential (ESP) of the amantadinium or memantinium cations. The molecular ESP of each cation was calculated at the MP2/6-31G (d,p) level of theory and the atomic charges were fitted using the Merz-Singh-Kollman scheme[232,233] as implemented in GAUSSIAN03. The Cl^- anion had a unit negative charge on the nucleus. The 1000 lowest energy structures were taken from each of the searches for **IX**

and **X** and passed to DMACRYS[89] for rigid-body lattice energy minimisations using a more accurate electrostatic model obtained by a distributed multipole analysis (DMA) of the MP2/6-31G (d,p) wavefunction. The DMA was performed using the program GDMA2.2[195] which allows for the multipoles to be calculated at foreshortened X-H positions (0.1 Å) as required for the W99 force field (see Table 2.1). Only structures at true lattice energy minima were kept. These structures were clustered[193] to remove multiple findings of the same lattice energy minima. This was done by comparing the relative lattice energies and cell densities, simulated X-ray powder diffraction patterns and COMPACK[115] overlays of the 15-molecule co-ordination spheres ($RMSD_{15}$) using a similarity threshold of 0.10 Å or less for identical structures. Molecular packing diagrams were created using the visualisation package, Mercury CSD 2.2[176]. All hydrogen bond interactions were defined as those where the donor-acceptor distances were less than the sum of the van der Waals radii of the constituent atoms. Bondi's compilation[234] of the van der Waals radii (1.70 Å for C, 1.20 Å for H, 1.55 Å for N and 1.75 Å for Cl) was used for the hydrogen bond definitions. For the predicted crystal structures of **IX** and **X**, PLATON[235] was used to calculate the percentage solvent accessible volume per unit cell. The default probe radius of 1.2 Å was used in the PLATON calculations. Mercury was used to visualise the calculated voids using a grid spacing of 0.2 Å and a probe radius of 1.2 Å. Not all predicted crystal structures were successfully passed through PLATON and in some instances the program failed to give an estimate of the percentage solvent accessible volume per unit cell.

6.3 Results and discussion

6.3.1 Crystal structures of amantadine hydrochloride

The known crystal structure (**FINVAZ**) of amantadine hydrochloride, **IX**, has a 1:1 molar ratio of the amantadinium and chloride ions in the asymmetric unit and the crystal belongs to the monoclinic space group, $C2/c$. Figure 6.1 shows the pleated ribbon [Graph set: $C_2^1(4)$] hydrogen bond motif of **FINVAZ**. In a CSD survey conducted as part of this thesis (Chapter 3, section 3.4.4), the $C_2^1(4)$ ribbon was found to be the most common hydrogen bond motif in chloride salts with an acyclic N^+ centre such as that found in the NH_3^+ groups of **IX** and **X**. The low temperature structure (*aman_103_0m*)

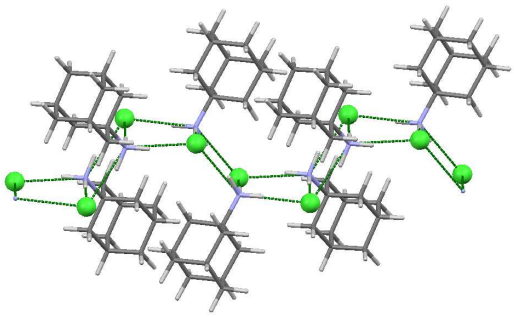
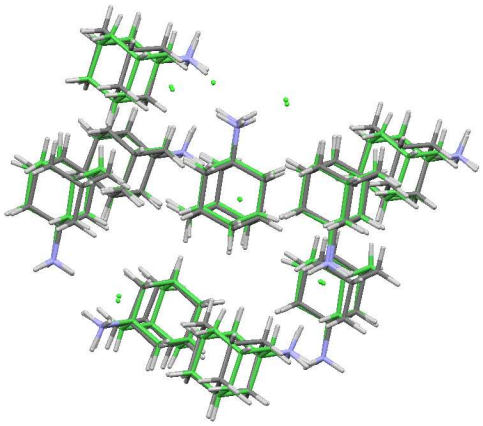
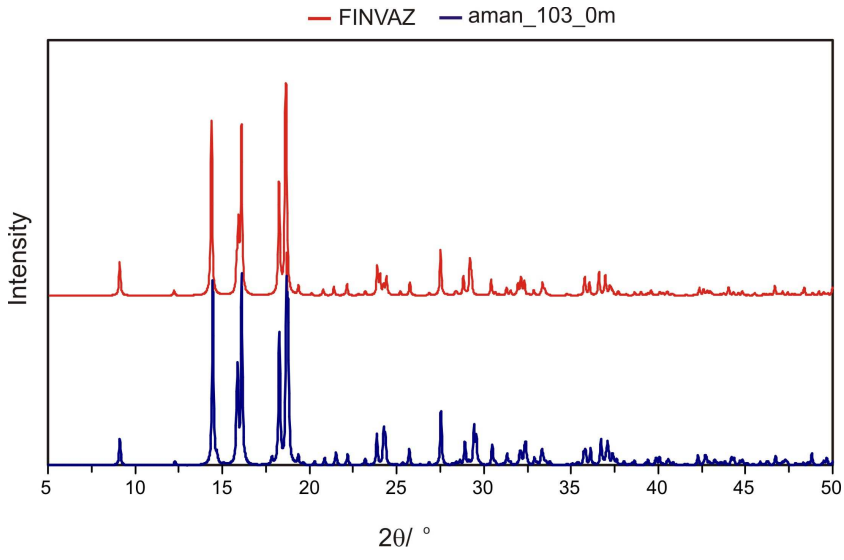


Figure 6.1: Illustration of the pleated ribbon hydrogen bond motif [graph set: $C_2^1(4)$] found in the structures of **FINVAZ** and *aman_103_0m*. The ribbons propagate parallel to the *c*-axis. The positions of the Cl^- anions are shown in bold green.



(a)



(b)

Figure 6.2: Shown in (a) is an overlay of the crystal packings for **FINVAZ** and *aman_103_0m* while (b) shows an overlay of the simulated X-ray powder diffraction patterns of the two structures assuming a wavelength of 1.54056 Å.

of **IX** crystallises in the same monoclinic space group as **FINVAZ** and has three cations and three anions in the crystallographic asymmetric unit. The *aman_103_0m* structure displays the same pleated ribbon hydrogen bond motif (Figure 6.1) found in **FINVAZ** and an overlay of the crystal packings (Figure 6.2) has shown that all 15 molecules of the co-ordination spheres match to give a root mean square deviation of 0.315 Å. An overlay of the simulated powder X-ray diffraction patterns (Figure 6.2) of the two structures suggests very little difference in the extended 3D packings. The similarity in the crystal packings of the two structures and the absence of major structural changes at the transition temperature is consistent with the reported[227] low enthalpy and entropy of transition at 124 K: ΔH = 0.31 kJ mol^{-1}, ΔS =2.54 J K^{-1} mol^{-1}. The small structural differences between **FINVAZ** and *aman_103_0m* can be explained if one regards the positions of the **FINVAZ** ions as a thermal average of three positions derived from the application of the molecular C_3 axis of the cation. In this mechanism, cooling a crystal of **FINVAZ** below 124 K freezes out the motions of the ions such that there are three crystallographically distinct positions for the cation and the anion. The evidence in support of this mechanism can be found in the high thermal parameters for atoms C3, C4, C8 and C10 of **FINVAZ**. According to the crystallographers who reported[226] the **FINVAZ** structure, the high thermal parameters are 'consistent with a slight oscillation or disorder about the C_3 axis of the molecule'.

6.3.2 Crystal structure prediction of amantadine hydrochloride

The predicted crystal energy landscape of **IX** is shown in Figure 6.3 and Table 6.1 (see section 6.5.1) provides structural parameters and packing motifs for the most stable lattice energy minima of **IX**. Because the crystal structure prediction assumed a 1:1 molar ratio of the amantadinium and chloride ions, only the **FINVAZ** structure was found in the search. The experimental **FINVAZ** structure is predicted to be the most stable at 0 K. There is an orthorhombic, *Pbcn* crystal based on the $C_2^1(4)$ pleated ribbon motif and this is approximately 0.98 kJ mol^{-1} less stable than the **FINVAZ** structure. The similarity in the crystal packings of the two structures is reflected in a root mean square deviation of 0.062 Å following a COMPACK overlay of the structures using a co-ordination sphere of 15 molecules ($RMSD_{15}$). However, when both structures are viewed down the *b*-axis, it is clear (Figure 6.4) that the ribbons have a

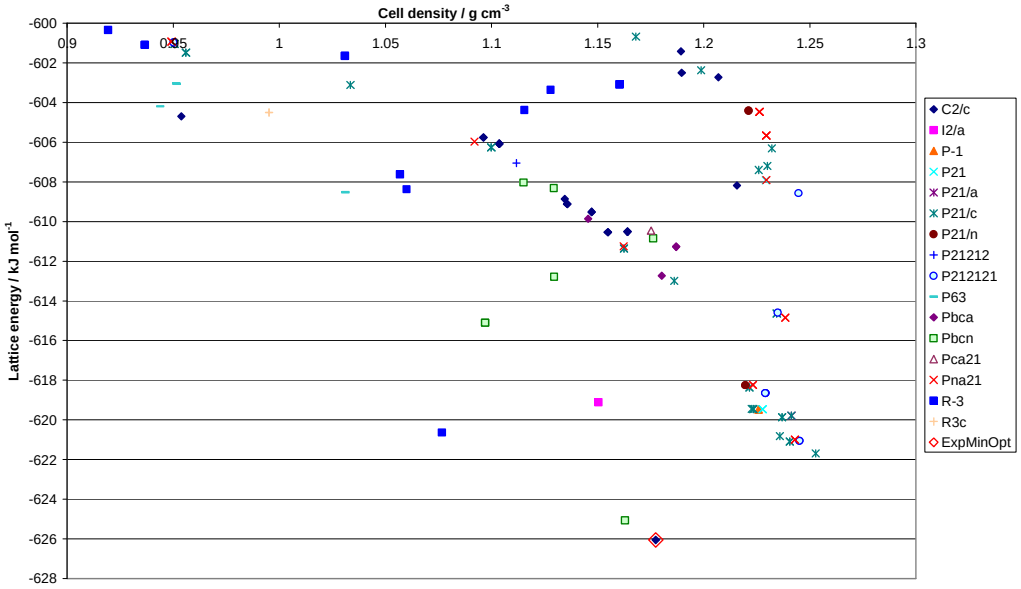
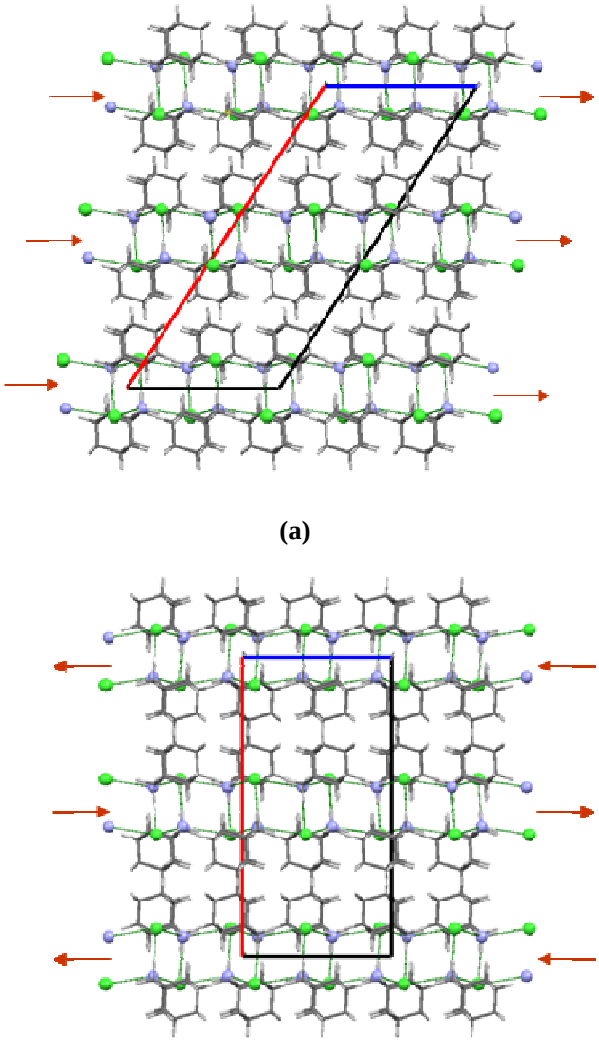


Figure 6.3: Predicted crystal energy landscape of amantadine hydrochloride, **IX**. All structures are denoted according to the space group.

parallel arrangement in the experimental structure and an anti-parallel arrangement in the second most stable structure. A phase transformation between the two structures would require a significant rearrangement of pleated ribbons and this is unlikely to be achieved by the application of shear stress alone. Direct nucleation is a more plausible mechanism for the crystallisation of a polymorph based on the anti-parallel arrangement of ribbons.

There is a 3.40 kJ mol^{-1} energy gap between the lowest energy orthorhombic structure and the next most stable structure in the predicted crystal energy landscape. Analysis of the hydrogen bond motifs adopted by all structures in Figure 6.3 showed that 75 % of the lattice energy minima were based on packings with one or two N^{+}-H donors that do not participate in hydrogen bond interactions as defined in section 6.2.1. In the majority of the cases, this was a consequence of a donor-acceptor distance that was longer than the sum of the van der Waals radii of the N and the Cl atoms. Figure 6.5 provides a visual comparison of some of the hydrogen bond motifs found among the most stable hypothetical crystal structures of **IX**. Of the 75 % of structures with one or more N^{+}-H donors not involved in conventional hydrogen bond interactions, 53 % (40 % of all lattice energy minima in Figure 6.3) of them were based on hydrogen bonded tapes that resemble the pleated ribbon motif of **FINVAZ.** A related 2D analogue of the

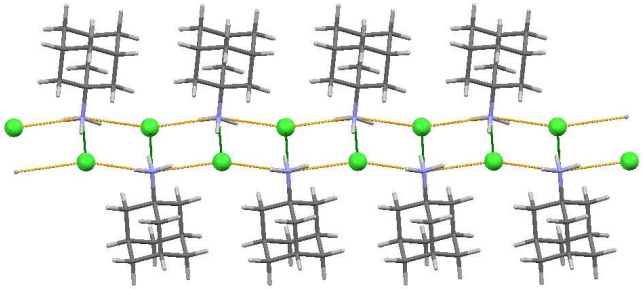


(a)

(b)

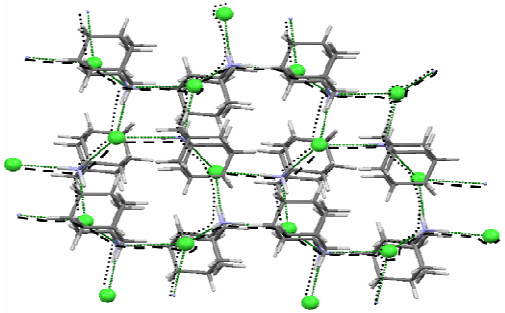
Figure 6.4: Packing of the $C_2^1(4)$ pleated ribbons found in the experimental (a) and the second most stable crystal structure (b) found in the crystal energy landscape of **IX**. In both diagrams, the packing is shown as viewed down the *b*-axis. The orange arrows indicate the direction of ribbon propagation.

experimental ribbon motif is also found among a handful of predicted structures which are energetically unfavourable. The fact that the majority of the low energy structures are based on hydrogen bonded tapes that resemble the experimentally observed ribbon motif, suggests that **IX** will have difficulties crystallising in a polymorphic modification with a drastically different hydrogen bonding motif to that found in **FINVAZ**. This observation was later confirmed by the finding of the structure (*aman_103_0m*) for the low temperature phase of **IX**, which consists of the $C_2^1(4)$ pleated ribbon hydrogen bond motif with the same parallel ribbon arrangement as found in **FINVAZ**. It must be remembered that although some of the motifs in Figure 6.5 are depicted as having



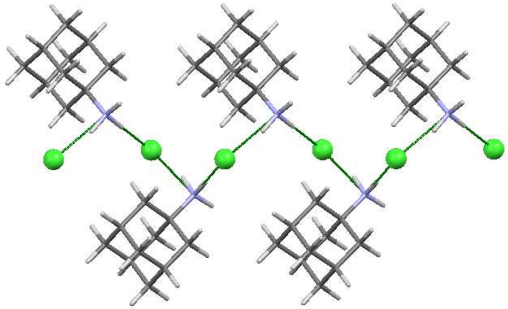
$D_1^1(2)$ – *TAPES*

2 unused donors per cation, 40 % occurrence



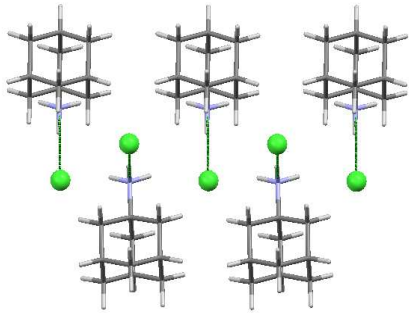
$C_2^1(4)$ – *2D RIBBONS*

All donors satisfied, 7 % occurrence



$C_2^1(4)$ – *1D CHAINS*

1 unused donor per cation, 7 % occurrence



$D_1^1(2)$ - *NO 1D/2D PATTERN*

2 unused donors per cation, 17 % occurrence

Figure 6.5: Illustration of some of the hydrogen bond motifs found among the low energy predicted crystal structures of **IX**. For comparison, the pleated ribbon motif of **FINVAZ** can be found in Figure 6.1. Hydrogen bond interactions are shown in green while interactions in orange are those $N^+\cdots Cl$ contacts that are within +0.2 Å of the sum of the van der Waals radii of the N and Cl atoms. For the 2D ribbon motif, the two directions of ribbon propagation are highlighted as the light/bold dashed lines. The quoted percentage occurrences take into account all the structures within 26 kJ mol^{-1} of the global minimum structure.

'unused' hydrogen bond donors (see section 6.2.1 for the definition of a hydrogen bond), the electrostatic interactions are long-range in nature. Thus even if the donor-acceptor distances are slightly longer (+0.2 Å for the tape motif) than the sum of the van der Waals radii for the constituent atoms, the interaction may still be strong.

The predicted crystal energy landscape (Figure 6.3) shows that there is a trigonal crystal structure ($R\bar{3}$) that is less dense but only 5.42 kJ mol^{-1} less stable than the experimental global minimum structure. This trigonal structure is based on hydrogen bonded hexagonal rings of graph set $R_6^3(12)$ and unlike the majority of the predicted structures shown in Figure 6.3, displays donor-acceptor distances that are within the sum of the van der Waals radii of the N and Cl atoms. This alternative $R_6^3(12)$ hydrogen bond motif is described in more detail below as it occurs in the crystal structure of memantine hydrochloride, **X**.

6.3.3 Anhydrate and monohydrate structures of memantine hydrochloride

The anhydrate crystal structure (*tinr3c0m*) of memantine hydrochloride, **X**, crystallises with a 1:1 molar ratio of the memantinium and chloride ions in the crystallographic asymmetric unit (Figure 6.6). The experimental structure is reported to have solvent accessible voids with a total volume of 240.7 Å^3 per unit cell. Details of the method used to calculate these voids were not given in the reported crystallographic information file. The residual electron density found in these voids could not be satisfactorily modelled by the group of Prof. Alistair Florence and instead they used the SQUEEZE[236] program to remove the equivalent of 12 electrons from the unit cell. When the crystal packing is viewed down the *c*-axis, the most intricate hydrogen bond motif consists of hexagonal shaped rings of graph set $R_6^3(12)$ formed from three memantinium and three chloride ions. The hexagonal rings only appear to be flat when viewed down the *c*-axis, but are actually puckered as a consequence of a $Cl^-\cdots N^+\cdots Cl$ angle of 110.43 °. Because of this, the actual shape of each $R_6^3(12)$ hexagonal ring resembles a chair conformation. Successive layers of the $R_6^3(12)$ hexagonal rings are related by a 3-fold rotation and this facilitates the formation of N^+-H$\cdots$Cl hydrogen bonds between these rings. Figure 6.7 depicts the crystal packing of *tinr3c0m* using a space filling model and the figure clearly shows the infinite channels formed from the stacking of the

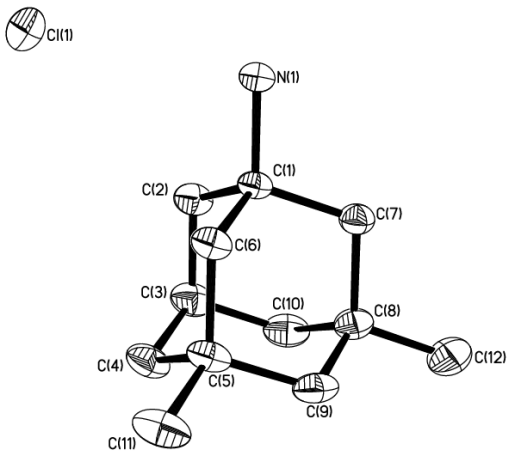


Figure 6.6: The asymmetric unit of anhydrate structure of memantine hydrochloride, **X**, found in the crystal structure of *tinr3c0m*. Displacement ellipsoids are drawn at the 50 % probability level and hydrogen atoms have been omitted for the sake of clarity.

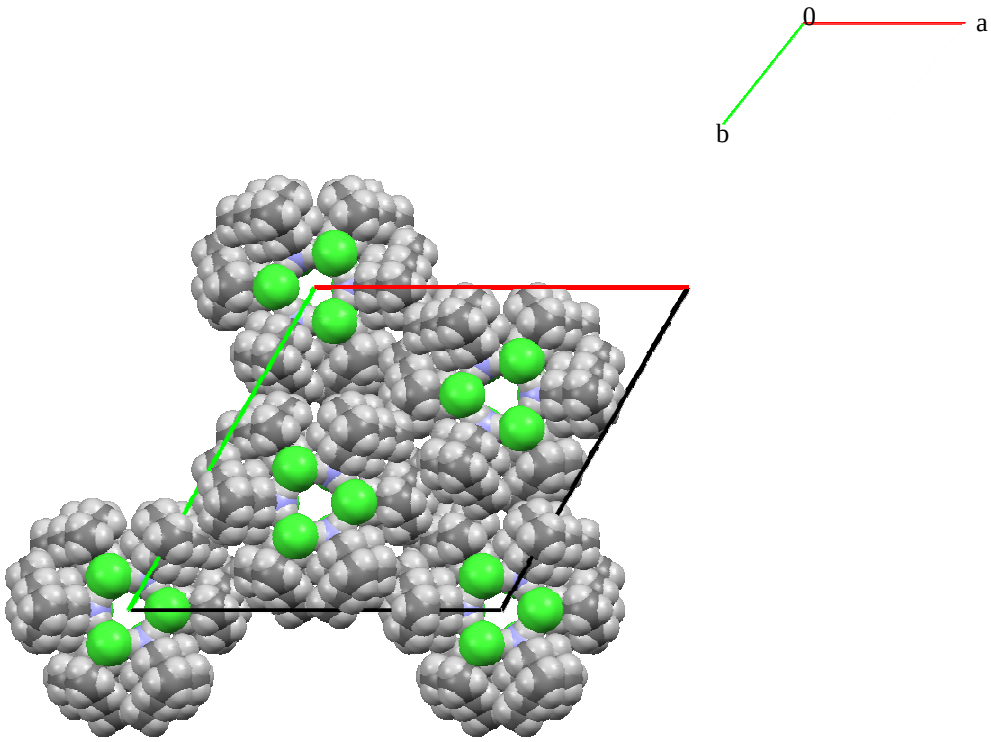


Figure 6.7: The crystal packing of the anhydrate structure of **X** (*tinr3c0m*) as viewed down the *c*-axis. All atoms are drawn using a space filling model. The regions of unoccupied volume between the chloride ions (coloured green) form channels that propagate parallel to the c-axis.

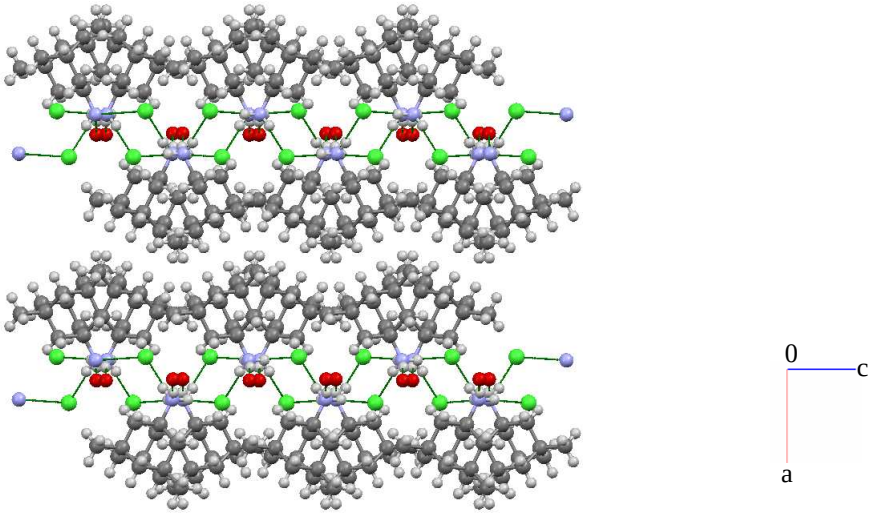


Figure 6.8: Crystal packing of the monohydrate structure of **X** as viewed down the b-axis. On each cation, two of the three N^+-H donors interact with a chloride anion [graph set: $C_2^1(4)$] while the third N^+-H donor forms a hydrogen bond with the oxygen of a water molecule.

$R_6^3(12)$ hexagonal rings. The channels have a diameter of 5.98 Å.

The monohydrate structure of **X** crystallises in the orthorhombic space group, *Pcab*. The positions of the water hydrogen atoms were not included in the refinement of the monohydrate crystal structure. Figure 6.8 shows the crystal packing of the monohydrate as viewed down the *b*-axis. The structure displays the same ribbons of N^+-H···Cl interactions [graph set: $C_2^1(4)$] found in the anhydrate structure. In the crystal structure of the monohydrate, only two of the three N^+-H donors on the cation are hydrogen bonded to a chloride anion. The third N^+-H donor is hydrogen bonded to the oxygen of a water molecule. Each water molecule acts as a bridge between adjacent layers of the ribbon motif and participates in hydrogen bond interactions with two Cl^- ions as well as an N^+-H donor of the cation.

6.3.4 Crystal structure prediction of memantine hydrochloride

The experimental structure (*tinr3c0m*) of memantine hydrochloride was successfully found in the predicted crystal energy landscape of the salt (Figure 6.9). A list of the structural parameters and hydrogen bond motifs of the most stable predicted lattice energy minima of **X** can be found in Table 6.2 of the appendix (section 6.5.2). The

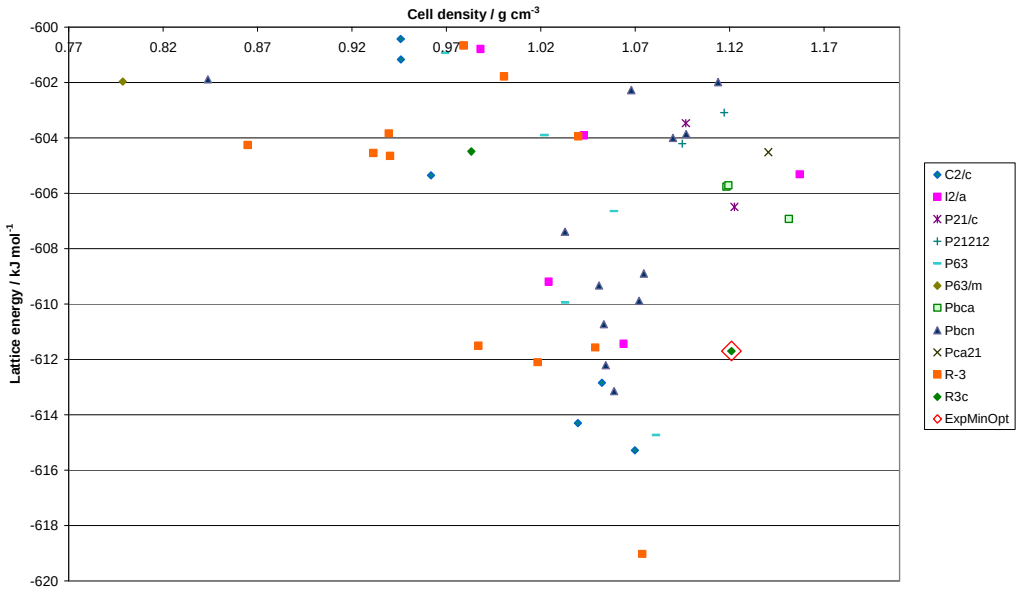
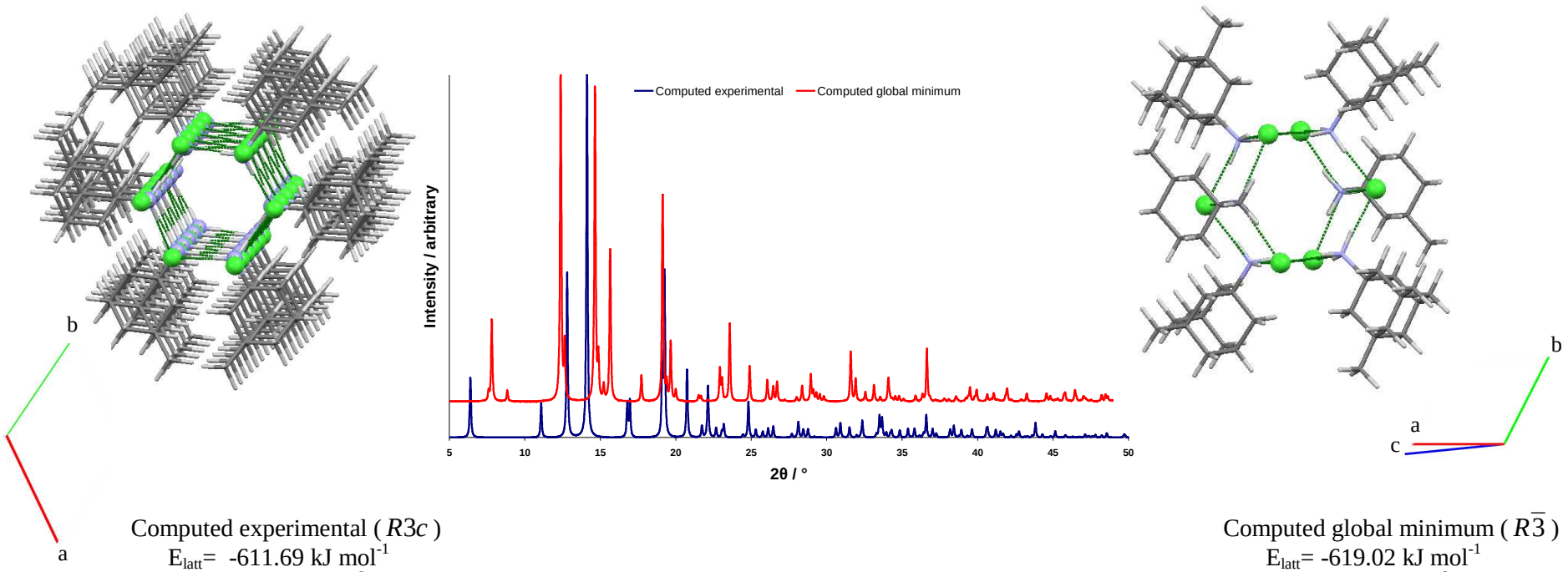


Figure 6.9: Scatter plot of the lattice energy versus cell density for the lowest energy predicted structures of memantine hydrochloride, **X**. All structures are denoted by space group.

experimental structure is ranked 9th in stability and is approximately 7.33 kJ mol^{-1} less stable than the global minimum structure. There are a variety of low symmetry, less dense structures with energetically more stable packing arrangements than the experimental structure. These include three monoclinic structures of space group $C2/c$ and two orthorhombic structures of space group $Pbcn$, all of which display the $C_2^1(4)$ pleated ribbon motif with parallel or anti-parallel arrangement of ribbons. However, the global minimum structure has the same hexagonal shaped rings [graph set: $R_6^3(12)$] of hydrogen bonded N^+-H$\cdots$Cl$^-$ interactions found in the experimental structure. The difference (Figure 6.10) is that in the experimental structure, the $R_6^3(12)$ rings stack on top of one another leading to infinite channels that propagate parallel to the *c*-axis. In the predicted global minimum structure, only inversion related pairs of hexagonal rings are observed and there are no infinite channels. This difference in the extended 3D packing of the crystals is sufficient to cause discernable differences in the simulated X-ray powder diffraction patterns (Figure 6.10) of the experimental and global minimum structures.

Inspection of the hydrogen bond motifs in the calculated crystal energy landscape (Figure 6.9) shows that 35 % of the lattice energy minima are based on the $C_2^1(4)$



Computed experimental ($R3c$)
E_{latt}= -611.69 kJ mol^{-1}
ρ_{calc}= 1.12 g cm^{-3}

Computed global minimum ($R\bar{3}$)
E_{latt}= -619.02 kJ mol^{-1}
ρ_{calc}= 1.07 g cm^{-3}

Figure 6.10: Shown to the sides are the hydrogen bond motifs in the experimental (left) and predicted global minimum structures (right) of **X** and some details about the calculated lattice energies and cell densities of the structures. A comparison of the simulated X-ray powder diffraction patterns of the two structures is shown in the middle of the page. The simulated powder patterns were produced with Mercury CSD 2.2 using all atoms in the structures and assuming a wavelength of 1.54056 Å.

pleated ribbon motif (parallel or anti-parallel arrangement) and this is the most commonly found motif in the calculated crystal energy landscape of **X**. The $R_6^3(12)$ motif accounts for 13 % of structures when the hexagonal rings are stacked to form infinite channels. The same motif accounts for 23 % of all lattice energy minima when the hexagonal rings are arranged as discrete pairs related by an inversion centre. Unlike the results from the crystal energy landscape of **IX**, the majority (87 %) of the predicted crystal structures of **X** satisfy all donors and acceptors in conventional hydrogen bond interactions. The exceptions are shown in Table 6.2 as structures based on the $R_4^2(8)$ hydrogen bond motif. The finding of a collection of hypothetical structures with more stable packing arrangements than the experimental, coupled with the competing hydrogen bond motifs found among the low energy structures, suggests that careful manipulation of the crystallisation conditions may lead to the finding of further polymorphs of **X**. Alternatively, the channel structure of the anhydrous phase may be an important feature of the crystallisation mechanism that prevents the nucleation of other polymorphs with alternative hydrogen bond motifs.

6.3.5 Investigating the propensity for hydration in crystals of IX and X

The predicted crystal energy landscapes of **IX** and **X** contain structures with N^+-H donors that do not participate in hydrogen bond interactions with the chloride ions and visual inspection of the crystal packings for a handful of structures, reveals voids suggestive of the possibility of hydration. Indeed a monohydrate structure of **X** has been found as part of the experimental screens. The anhydrous crystal structure of **X** has solvent accessible voids within the channels formed by the $R_6^3(12)$ hydrogen bond motif (Figure 6.7). By contrast, no hydrates of **IX** are known. Given the different hydration behaviour of these salts, we decided to investigate the utility of the predicted crystal energy landscapes to help rationalise this difference.

In an attempt to understand the role played by water in the crystallisation of the monohydrate, a *computational dehydration experiment* was performed by removing the water molecule from the monohydrate and allowing the structure to relax under the packing forces. The resulting lattice energy minimum (denoted *DehydMinOpt*) increased its density by 4 % after lattice energy minimisation, having contracted in the

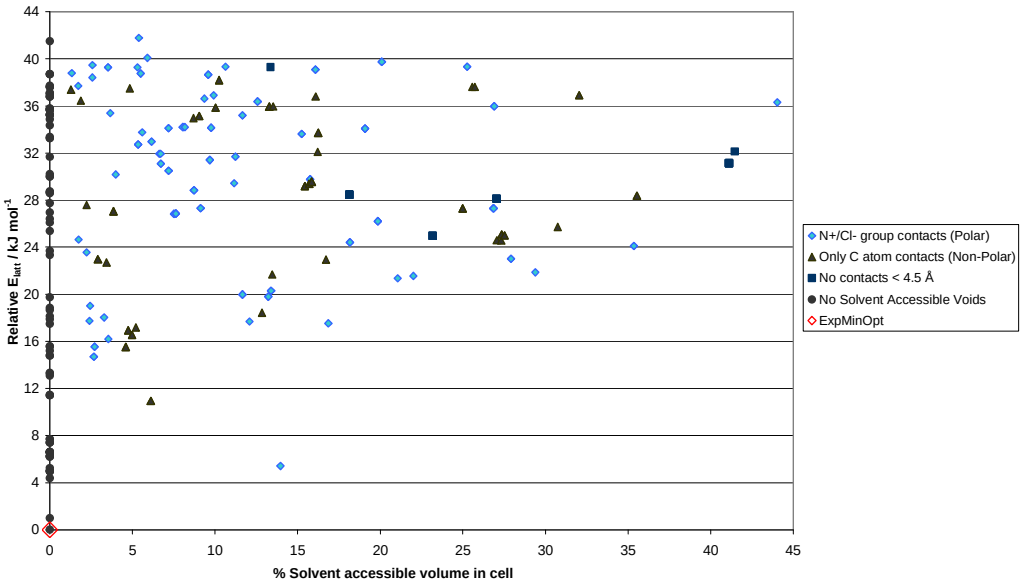


Figure 6.11: Plot of the % solvent accessible volume in the unit cell versus the relative lattice energy for the predicted* crystal structures of amantadine hydrochloride, **IX**.

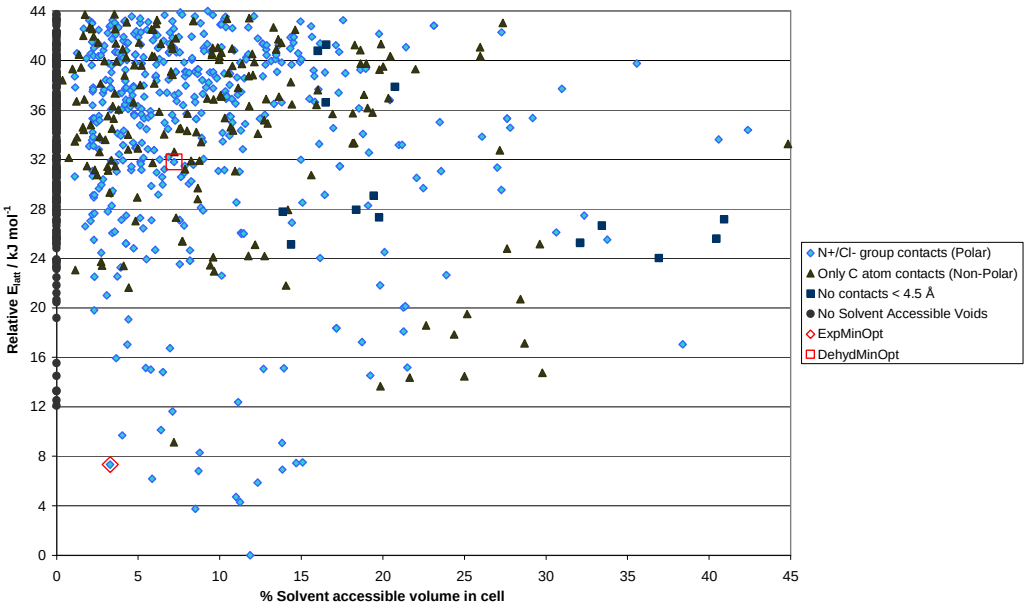


Figure 6.12: Plot of the % solvent accessible volume in the unit cell versus the relative lattice energy for the predicted* crystal structures of memantine hydrochloride, **X**.

* For a small proportion of structures, PLATON failed to give an estimate of the % solvent accessible volume in the cell.

ab plane (4 % along *a*, 8 % along *b*) but expanded along the *c*-axis. The ribbons of N^+-H···Cl interactions in the monohydrate (Figure 6.8) remain intact in the *DehydMinOpt* but are closer together as a consequence of the ions occupying the vacancies created by the water molecules. The *DehydMinOpt* was compared with the hypothetical structures in the crystal energy landscape of **X** and structure 573 found as a match (Figure 6.13). This structure displays a ribbon hydrogen bond motif [graph set: $C_2^1(4)$] with two of the three N^+-H protons on each cation participating in hydrogen bond interactions with a chloride ion and has a relative lattice energy of 24.47 kJ mol^{-1} with respect to the experimental structure. The third N^+-H proton is an unused hydrogen bond donor. Comparison of structure 573 with the monohydrate structure of **X** shows that the qualitative difference between the two structures is that in the monohydrate the water molecule stabilises this unused hydrogen bond donor via the formation of strong N^+-H···O_{water} interactions. Powder X-ray diffraction measurements into the hydration/dehydration mechanism in **X** have shown slow interconversion between the anhydrous trigonal structure and the orthorhombic monohydrate structure. The % solvent accessible volume in the *DehydMinOpt* is 7.2 % in contrast to 8.4 % in the experimental hydrate structure with just the water removed.

Recent work by Day *et al*[237], successfully correlating high energy/low density structures of hydroquinone and urea with their clathrate frameworks, has shown that careful analysis of the amount of void volume found in the predicted crystal structures can help rationalise the solvation behaviour of framework structures. PLATON[148] was used to calculate the amount of void volume in each of the hypothetical crystal structures of **IX** and **X**. Figure 6.11 and Figure 6.12 are plots of the % solvent accessible volume in the unit cell versus the relative lattice energy for the hypothetical structures in the crystal energy landscapes of **IX** and **X** respectively. The structures are denoted according to the short contacts found within 4.5 Å of the calculated void regions and all lattice energies are quoted relative to the predicted global minimum structure. For amantadine hydrochloride, **IX**, Figure 6.11 shows that the experimental structure has no solvent accessible voids. This suggests minimal risk of hydration and agrees with the results from experimental screens which have not uncovered a hydrate crystal form. Within a 12 kJ mol^{-1} range in the relative lattice energy, there are only two hypothetical structures of **IX** with solvent accessible voids. The first is an $R\bar{3}$ trigonal structure with the same hexagonal motif found in the predicted global minimum

structure of memantine hydrochloride, **X**. This structure has a relative lattice energy of 5.42 kJ mol^{-1} and contains voids that are located within 4.5 Å of the charged N^+/Cl^- centres. The second structure has a relative lattice energy of 10.95 kJ mol^{-1} and the only short contacts within 4.5 Å of the calculated voids are those with C atoms. For memantine hydrochloride, **X**, the majority of the predicted low energy structures, including the experimental *tinr3c0m* structure, have solvent accessible voids (Figure 6.12). This suggests some risk of hydration and the results agree with the experimental finding of a monohydrate structure of **X**. With the exception of one structure, all the predicted structures within 12 kJ mol^{-1} of the global minimum are characterised by voids that are located within 4.5 Å of the charged N^+/Cl^- centres. The fact that the majority of the predicted low energy structures of **X** have voids that are in close proximity to the N^+/Cl^- centres suggests that polar solvents like water are good candidates for forming solvates of memantine hydrochloride. The predicted experimental structure of **X** has the lowest % solvent accessible volume (3.28 % of unit cell volume) of all the structures within a 12 kJ mol^{-1} range in the relative lattice energy.

6.4 Conclusions

The experimental screen for the low temperature phase of amantadine hydrochloride, **IX**, led to the finding of a structure (*aman_103_0m*) with three cations and three anions in the crystallographic asymmetric unit. This structure is based on the same pleated ribbon hydrogen bond motif found in the known crystal structure of amantadine hydrochloride (**FINVAZ**). The Niggli reduced cells of the *aman_103_0m* and **FINVAZ** structures were found to be different. Computational modelling of the two structures has shown that they both relax to the same lattice energy minimum. Thus there are two distinct crystal structures of **IX** at 143 K (**FINVAZ**) and 105 K (*aman_103_0m*) and although the packing difference is subtle, the observed phase transition suggests that they are genuine polymorphs. Crystal structure prediction of a 1:1 molar ratio of the amantadinium and chloride ions has shown that the experimental **FINVAZ** structure of **IX** is the global minimum in the crystal energy landscape of the salt. The majority of the calculated low energy structures have N^+-H donors that do not participate in conventional hydrogen bond interactions as a consequence of donor-acceptor distances that are longer than the sum of the van der Waals radii of the N and

Cl atoms. The most commonly found hydrogen bond motif is based on discrete N^+-H···Cl pairs of cation and anion [graph set: $D_1^1(2)$] interactions, which are organised in the form of tapes resembling the pleated ribbon motif of **FINVAZ**. The absence of many competing hydrogen bond motifs among the low energy structures was interpreted as indicating difficulties in crystallising a polymorphic modification of **IX** with a drastically different packing to that found in **FINVAZ**. The finding of a low temperature structure of **IX** (*aman_103_0m*) based on the same pleated ribbon motif of **FINVAZ** later confirmed this. Analysis of the solvent accessible voids in the predicted lattice energy minima of **IX** has shown that the **FINVAZ** structure is unlikely to form a hydrate since it does not contain solvent accessible voids. After a range of crystallisation experiments, the lack of any hydrated crystal forms of **IX** indicated the predictive power of crystal energy landscapes in warning against the possibility of hydration.

For memantine hydrochloride, **X**, experiments at Strathclyde led to anhydrous (*tinr3c0m*) and monohydrate phases of the salt with significant difference in the packing of the ions. During the course of the theoretical calculations, other workers have published the *tinr3c0m* structure as a 10% hydrate. In the calculated crystal energy landscape of anhydrous memantine hydrochloride, the experimental structure is ranked 9th in stability and there are a handful of alternative low energy packing arrangements of **X** including some of lower symmetry based on the pleated ribbon motif of **FINVAZ**. Comparison of the most stable predicted structure of **X** based on this ribbon motif and the predicted experimental structure of **IX** has shown slight differences in the positions of the ions but the same overall motif including the direction of ribbon propagation. The calculated global minimum structure of **X** belongs to the trigonal space group $R\bar{3}$ and is 7.33 kJ mol^{-1} more stable than the experimental structure. This structure has the same hexagonal ring motif [graph set: $R_6^3(12)$] found in the experimental structure, but is packed in a way that does not result in the infinite channels found in *tinr3c0m*. Hence there are alternative structures that may be experimentally accessible polymorphs of memantine hydrochloride. Alternatively, the channels in the observed form may be crucial to the kinetic pathway to an anhydrous structure, thereby preventing the crystallisation of alternative polymorphs not based on the channel motif. A computational dehydration of the monohydrate has shown that the resulting lattice

energy minimum (*DehydMinOpt*) was predicted in the search but is 24.47 kJ mol^{-1} higher in energy than the experimental structure. Many of the predicted low energy structures of **X** - including the experimental structure - contain solvent accessible voids in contrast to the results for **IX**. Thus it seems likely that there could be other hydrates of **X**.

The success in predicting the known structure of **IX** as the thermodynamically most stable form is strong evidence that these chloride salt lattice energy landscapes are reasonably realistic. Within the usual caveats, the calculations suggest that the crystallisation of amantadine hydrochloride is thermodynamically controlled, and it has a significantly favoured means of packing with itself. In contrast, the ability to pack the same hydrogen bonded motifs with the extra methyl groups in **X** result in more complex crystallisation behaviour. The results from the crystal energy landscape suggest that the anhydrous structure of **X** already determined may be a kinetic rather than thermodynamic product, at least at low temperature.

6.5 Appendix

6.5.1 Predicted most stable lattice energy minima of amantadine hydrochloride IX

Structure ID	E_{latt} / kJ mol^{-1}	ρ / g cm^{-3}	Space Group	Hydrogen bond motif	a / Å	b / Å	c / Å	α / °	β / °	γ / °
FINVAZ	-	1.191	C2/c	$C_2^1(4)$ - PLEATED RIBBON	20.549	11.138	9.658	90	108.810	90
395	*-626.055*	*1.177*	*C2/c*	$C_2^1(4)$ - *PLEATED RIBBON*	*23.250*	*11.077*	*9.823*	*90*	*123.159*	*90*
5	-625.069	1.163	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	19.603	11.227	9.743	90	90	90
8	-621.683	1.253	$P2_1/c$	$D_1^1(2)$ - TAPE	11.372	6.297	16.027	90	119.868	90
16	-621.098	1.241	$P2_1/c$	$D_1^1(2)$ - TAPE	6.321	15.539	11.797	90	60.155	90
10	-621.053	1.245	$P2_12_12_1$	$D_1^1(2)$ - TAPE	10.196	6.310	15.563	90	90	90
14	-620.998	1.243	$Pna2_1$	$D_1^1(2)$ - TAPE	15.527	10.220	6.321	90	90	90
29	-620.820	1.236	$P2_1/c$	$D_1^1(2)$ - TAPE	6.351	10.227	15.532	90	89.9999 7	90
38	-620.638	1.077	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	14.204	14.204	29.825	90	90	120
22	-619.858	1.237	$P2_1/c$	$D_1^1(2)$ - TAPE	7.975	6.318	20.657	90	104.443	90
1156	-619.790	1.241	$P2_1/a$	$D_1^1(2)$ - TAPE	15.954	6.303	16.835	90	36.391	90
33	-619.476	1.226	P-1	$D_1^1(2)$ - TAPE	7.973	6.341	10.384	91.546	104.306	89.64713
41	-619.467	1.228	$P2_1$	$D_1^1(2)$ - TAPE	10.379	6.336	7.988	90	104.855	90
48	-619.445	1.223	$P2_1/c$	$D_1^1(2)$ - TAPE	7.978	6.353	20.789	90	75.285	90
54	-619.444	1.223	$P2_1/c$	$D_1^1(2)$ - TAPE	11.371	6.359	15.959	90	117.906	90
28	-619.113	1.150	I2/a	$C_2^1(4)$ - PLEATED RIBBON	9.787	21.762	10.302	90	81.102	90
30	-618.652	1.229	$P2_12_12_1$	$D_1^1(2)$ - TAPE	20.259	7.947	6.301	90	90	90

43	-618.366	1.222	$P2_1/c$	$R_4^2(8)$ - TAPE	6.320	7.911	20.425	90	91.736	90
49	-618.246	1.220	$P2_1/n$	$D_1^1(2)$ - TAPE	6.326	20.505	7.882	90	89.031	90
45	-618.242	1.223	$Pna2_1$	$D_1^1(2)$ - TAPE	20.397	7.903	6.323	90	90	90
75	-615.104	1.097	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	10.987	21.018	9.843	90	90	90
64	-614.845	1.238	$Pna2_1$	$D_1^1(2)$ - TAPE	13.676	11.661	6.313	90	90	90
71	-614.644	1.234	$P2_1/c$	$D_1^1(2)$ - TAPE	6.350	13.719	13.218	90	118.705	90
70	-614.594	1.235	$P2_12_12_1$	$D_1^1(2)$ - TAPE	13.715	11.571	6.362	90	90	90
81	-612.981	1.186	$P2_1/c$	$R_4^2(8)$	11.263	9.659	9.720	90	83.717	90
108	-612.790	1.129	Pbcn	$R_4^2(8)$	10.663	20.154	10.273	90	90	90
77	-612.735	1.180	Pbca	$C_2^1(4)$ - CHAIN	9.806	9.087	23.710	90	90	90
98	-611.365	1.162	$P2_1/c$	$D_1^1(2)$ - TAPE	6.308	21.465	7.924	90	91.492	90
84	-611.270	1.187	Pbca	$D_1^1(2)$	22.426	9.681	9.676	90	90	90
106	-611.241	1.162	$Pna2_1$	$D_1^1(2)$ - TAPE	7.929	21.338	6.340	90	90	90
87	-610.851	1.176	Pbcn	$D_1^1(2)$	23.892	9.726	9.124	90	90	90
101	-610.538	1.155	C2/c	$D_1^1(2)$ - TAPE	24.703	6.290	16.072	90	59.854	90
96	-610.508	1.164	C2/c	$D_1^1(2)$	25.048	9.621	9.389	90	71.223	90
89	-610.454	1.175	$Pca2_1$	$D_1^1(2)$	9.730	11.952	9.123	90	90	90
1056	-609.856	1.145	Pbca	$C_2^1(4)$ / $R_4^2(8)$ - 2D RIBBON	9.329	23.903	9.762	90	90	90
111	-609.509	1.147	C2/c	$D_1^1(2)$ - TAPE	21.918	6.289	15.988	90	99.483	90
120	-609.115	1.136	C2/c	$D_1^1(2)$ - TAPE	22.103	6.293	16.060	90	79.389	90
141	-608.868	1.135	C2/c	$D_1^1(2)$ - TAPE	24.937	6.344	15.985	90	60.362	90
114	-608.564	1.245	$P2_12_12_1$	$C_2^1(4)$ - CHAIN	21.448	6.398	7.300	90	90	90

118	-608.521	1.031	$P6_3$	$R_6^3(12)$ - CHANNELS	15.533	15.533	8.680	90	90	120
147	-608.369	1.060	$R\bar{3}$	$R_6^3(12)$ / $R_4^2(8)$	12.718	12.718	37.788	90	90	120
131	-608.314	1.129	Pbcn	$C_2^1(4)$ / $R_4^2(8)$ - 2D RIBBON	10.128	22.720	9.596	90	90	90
155	-608.176	1.216	C2/c	$R_4^2(8)$	20.965	9.965	10.217	90	106.049	90
142	-608.024	1.115	Pbcn	$C_2^1(4)$ / $R_4^2(8)$ - 2D RIBBON	9.970	22.689	9.885	90	90	90
128	-607.928	1.229	$P2_1$	$C_2^1(4)$ - CHAIN	6.394	7.203	11.041	90	94.223	90
122	-607.897	1.230	$Pna2_1$	$C_2^1(4)$ - CHAIN	22.062	6.382	7.202	90	90	90
152	-607.619	1.057	$R\bar{3}$	$R_6^3(12)$ / $R_4^2(8)$	12.706	12.706	37.967	90	90	120
137	-607.399	1.226	$P2_1/c$	$C_2^1(4)$ - CHAIN	6.349	7.232	22.208	90	94.062	90
127	-607.197	1.230	$P2_1/c$	$D_1^1(2)$	11.258	7.236	12.444	90	90.149	90
161	-607.050	1.112	$P2_12_12$	$C_2^1(4)$ / $R_4^2(8)$ - 2D RIBBON	9.975	9.717	11.571	90	90	90
136	-606.315	1.232	$P2_1/c$	$D_1^1(2)$	12.791	7.230	12.460	90	118.566	90
168	-606.262	1.100	$P2_1/c$	$D_1^1(2)$ - TAPE	6.315	17.726	10.161	90	94.709	90
163	-606.074	1.104	C2/c	$D_1^1(2)$ - TAPE	18.034	6.317	20.190	90	79.191	90
211	-605.960	1.092	$Pna2_1$	$D_1^1(2)$ - TAPE	10.154	17.673	6.363	90	90	90
183	-605.762	1.096	C2/c	$D_1^1(2)$ - TAPE	17.991	6.347	20.274	90	79.317	90
157	-605.664	1.230	$Pna2_1$	$D_1^1(2)$	22.210	6.378	7.158	90	90	90
782	-604.696	0.954	C2/c	$C_2^1(4)$ - PLEATED RIBBON	10.553	26.617	9.410	90	81.540	90
195	-604.501	0.995	R3c	$R_6^3(12)$ - CHANNELS	27.489	27.489	8.615	90	90	120
181	-604.473	1.226	$Pna2_1$	$D_1^1(2)$	22.204	7.214	6.348	90	90	90
199	-604.404	1.221	$P2_1/n$	$D_1^1(2)$	7.203	22.312	6.354	90	88.997	90

177	-604.379	1.115	$R\bar{3}$	$D_1^1(2)$	13.723	13.723	30.842	90	90	120
206	-604.187	0.944	$P6_3$	$R_6^3(12)$ - CHANNELS	16.238	16.238	8.678	90	90	120
173	-603.351	1.128	$R\bar{3}$	$R_4^2(8)$	22.416	22.416	11.432	90	90	120
226	-603.102	1.034	$P2_1/c$	$C_2^1(4)$ / $R_4^2(8)$ - 2D RIBBON	13.619	9.796	9.344	90	75.400	90
228	-603.085	1.160	$R\bar{3}$	$D_1^1(2)$	20.967	20.967	12.701	90	90	120
240	-603.041	0.951	$P6_3$	$R_6^3(12)$ - CHANNELS	16.255	16.255	8.590	90	90	120
289	-602.722	1.207	C2/c	$R_4^2(8)$	10.185	10.051	20.188	90	88.566	90
238	-602.501	1.190	C2/c	$D_1^1(2)$	24.157	7.789	12.734	90	118.978	90
248	-602.368	1.199	$P2_1/c$	$R_4^2(8)$	7.888	12.627	11.376	90	113.378	90
258	-601.643	1.031	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	25.993	25.993	9.301	90	90	120
279	-601.479	0.956	$P2_1/c$	$D_1^1(2)$	6.318	11.985	20.643	90	56.544	90
244	-601.416	1.189	C2/c	$D_1^1(2)$	21.278	7.743	12.735	90	87.913	90
252	-601.093	0.937	$R\bar{3}$	$R_4^2(8)$	33.214	33.214	6.270	90	90	120
292	-601.045	0.951	$P2_1/c$	$D_1^1(2)$ - TAPE	6.348	17.412	13.361	90	62.638	90
300	-600.977	0.950	$P2_12_12_1$	$D_1^1(2)$ - TAPE	17.418	6.358	11.850	90	90	90
315	-600.927	0.949	$Pna2_1$	$D_1^1(2)$ - TAPE	17.516	11.759	6.377	90	90	90
284	-600.675	1.168	$P2_1/c$	$R_4^2(8)$	9.753	10.320	12.460	90	58.335	90
335	-600.337	0.919	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	20.739	20.739	16.384	90	90	120

Table 6.1: Predicted low energy crystal structures of amantadine hydrochloride, **IX**. Structure 395 (highlighted in italics) is the predicted global minimum structure corresponding to **FINVAZ**. The structural parameters for **FINVAZ** are shown at the top of the table for comparison with structure 395.

6.5.2 Predicted most stable lattice energy minima of memantine hydrochloride X

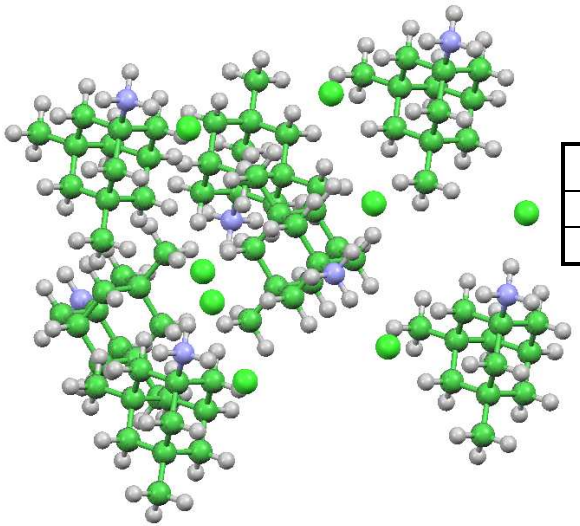
STRUCTURE ID	E_{latt} / kJ mol^{-1}	ρ / g cm^{-3}	Space Group	Hydrogen bond motif	a / Å	b / Å	c / Å	α / °	β / °	γ / °
tinr3c0m	-	1.115	R3c	$R_6^3(12)$ - CHANNELS	28.164	28.164	8.420	90	90	120
1900	-619.023	1.074	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	14.294	14.294	33.940	90	90	120
3	-615.280	1.070	C2/c	$C_2^1(4)$ - PLEATED RIBBON	21.244	13.690	10.214	90	64.415	90
2	-614.728	1.081	$P6_3$	$R_6^3(12)$ - CHANNELS	16.290	16.290	8.653	90	90	120
4	-614.303	1.040	C2/c	$C_2^1(4)$ - PLEATED RIBBON	22.003	13.771	9.717	90	69.457	90
8	-613.148	1.059	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	19.117	13.630	10.389	90	90	90
7	-612.840	1.052	C2/c	$C_2^1(4)$ - PLEATED RIBBON	23.926	12.389	10.098	90	65.494	90
6	-612.209	1.054	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	20.326	13.384	9.993	90	90	90
13	-612.097	1.018	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	16.307	16.307	27.500	90	90	120
5	*-611.694*	*1.121*	*R3c*	$R_6^3(12)$ - *CHANNELS*	*27.641*	*27.641*	*8.695*	*90*	*90*	*120*
9	-611.568	1.049	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	14.274	14.274	34.842	90	90	120
16	-611.507	0.987	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	16.929	16.929	26.327	90	90	120
11	-611.435	1.064	I2/a	$C_2^1(4)$ - PLEATED RIBBON	9.533	24.918	11.366	90	86.304	90
193	-610.727	1.053	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	11.065	26.121	9.414	90	90	90
10	-609.942	1.033	$P6_3$	$R_6^3(12)$ - CHANNELS	16.740	16.740	8.576	90	90	120
17	-609.883	1.072	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	10.900	23.961	10.237	90	90	90
15	-609.331	1.051	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	21.733	12.159	10.322	90	90	90
12	-609.195	1.024	I2/a	$C_2^1(4)$ - PLEATED RIBBON	9.581	23.648	12.352	90	89.508	90
19	-608.899	1.075	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	17.624	13.008	11.634	90	90	90
21	-607.393	1.033	Pbcn	$R_4^2(8)$	15.828	14.965	11.716	90	90	90

18	-606.929	1.151	Pbca	$C_2^1(4)$ - PLEATED SHEET	10.895	8.770	26.054	90	90	90
22	-606.647	1.059	$P6_3$	$R_6^3(12)$ - CHANNELS	16.437	16.437	8.678	90	90	120
20	-606.491	1.123	$P2_1/c$	$C_2^1(4)$ / $R_4^2(8)$ - 2D RIBBON	13.036	10.887	8.996	90	89.664	90
24	-605.759	1.118	Pbca	$C_2^1(4)$ / $R_4^2(8)$ - 2D RIBBON	26.019	10.977	8.974	90	90	90
504	-605.715	1.119	Pbca	$C_2^1(4)$ / $R_4^2(8)$ - 2D RIBBON	8.990	25.998	10.957	90	90	90
23	-605.357	0.962	C2/c	$C_2^1(4)$ - PLEATED RIBBON	17.723	18.788	9.966	90	63.892	90
27	-605.321	1.157	I2/a	$R_4^2(8)$	10.265	21.962	11.072	90	82.949	90
47	-604.651	0.940	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	17.991	17.991	24.472	90	90	120
38	-604.547	0.931	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	20.231	20.231	19.536	90	90	120
26	-604.518	1.141	$Pca2_1$	$C_2^1(4)$ - PLEATED SHEET	10.947	13.239	8.670	90	90	90
28	-604.491	0.983	R3c	$R_6^3(12)$ - CHANNELS	29.669	29.669	8.604	90	90	120
39	-604.265	0.865	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	20.933	20.933	19.651	90	90	120
1051	-604.210	1.095	$P2_12_12$	$R_4^2(8)$	10.434	9.559	13.124	90	90	90
33	-604.008	1.090	Pbcn	$R_4^2(8)$	9.592	26.350	10.404	90	90	90
37	-603.951	1.040	$R\bar{3}$	$R_6^3(12)$ / $C_2^1(4)$ / $R_4^2(8)$	13.404	13.404	39.857	90	90	120
35	-603.911	1.043	I2/a	$C_2^1(4)$ - PLEATED RIBBON	9.847	27.451	10.182	90	87.160	90
31	-603.900	1.022	$P6_3$	$R_6^3(12)$ - CHANNELS	16.762	16.762	8.645	90	90	120
29	-603.865	1.097	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	16.731	12.979	12.031	90	90	90
42	-603.841	0.940	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	17.884	17.884	24.781	90	90	120
515	-603.473	1.097	$P2_1/c$	$C_2^1(4)$ / $R_4^2(8)$ - 2D RIBBON	15.462	10.856	9.038	90	59.462	90
32	-603.095	1.117	$P2_12_12$	$R_4^2(8)$	10.335	10.335	12.011	90	90	90
45	-602.275	1.068	Pbcn	$R_4^2(8)$	10.180	26.705	9.873	90	90	90
40	-601.987	1.114	Pbcn	$R_4^2(8)$	10.351	24.190	10.278	90	90	90

55	-601.967	0.799	$P6_3/m$	$R_6^3(12)$ - DISCRETE RINGS	14.202	14.202	30.822	90	90	120
49	-601.889	0.844	Pbcn	$C_2^1(4)$ - PLEATED RIBBON	16.507	22.071	9.324	90	90	90
66	-601.784	1.001	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	29.218	29.218	8.719	90	90	120
51	-601.165	0.946	C2/c	$C_2^1(4)$ - PLEATED RIBBON	10.773	29.776	9.465	90	93.640	90
48	-600.930	0.969	$P6_3$	$R_6^3(12)$ - CHANNELS	17.091	17.091	8.766	90	90	120
53	-600.798	0.988	I2/a	$C_2^1(4)$ - PLEATED RIBBON	10.136	25.309	11.393	90	82.994	90
128	-600.674	0.979	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	25.998	25.998	11.251	90	90	120
56	-600.663	0.979	$R\bar{3}$	$R_6^3(12)$ - DISCRETE RINGS	25.999	25.999	11.251	90	90	120
254	-600.434	0.946	C2/c	$C_2^1(4)$ - PLEATED RIBBON	11.076	27.142	10.082	90	89.125	90

Table 6.2: Predicted low energy crystal structures of memantine hydrochloride, **X**. Structures 5 and 1900 are the predicted experimental (*tinr3c0m*) and global minimum structures respectively. The lattice parameters for the experimental *tinr3c0m* structure are shown at the top of the table for comparison with structure 5 (highlighted in italics).

6.5.3 Comparison of the DehydMinOpt and structure 573 from the crystal energy landscape of memantine hydrochloride X



Structure	E_{latt} / kJ mol^{-1}	Space group	a / Å	b / Å	c / Å	α / °	β / °	γ / °
DehydMinOpt	-587.2254	Pcab	25.3773	11.5307	9.0061	90	90	90
573	-587.2280	Pbca	9.0061	11.5307	25.3774	90	90	90

Figure 6.13: Shown on the left is an overlay of the crystal packings for the lattice energy minimum calculated following relaxing the monohydrate structure of **X** without the water (*DehydMinOpt*) and structure 573 from the crystal structure prediction of anhydrous **X**. The two structures match as indicated by the overlay of the *DehydMinOpt* (green) and structure 573 (coloured by element) packings shown on the left ($RMSD_{15}$=0 Å). Details of the lattice energies and cell parameters of the two structures are shown on the right. Note that the *DehydMinOpt* and structure 573 have alternative cell settings of space group number 61.

6.5.4 Validating the intermolecular dispersion-repulsion model

Extensive[77,238,239] use has been made of empirically derived isotropic atom-atom potentials for modelling the dispersion-repulsion contributions towards the lattice energy of crystal structures consisting of neutral organic molecules. For ionic systems, there are many examples where empirically derived isotropic atom-atom potentials have been successfully applied in simulation work, but the use of such potential models is more common among simple alkali or ammonium salts[240-242] than they are among the structures of molecular[131,135,243] organic salts. The term "molecular organic salt" has been used by Nangia[156] to refer to salts where both the cation and anion are derived from neutral organic molecules, but here it is used to refer to any salt where at least one of the ions in the salt crystal structure is derived from a neutral organic molecule. For molecular organic salts, dispersion-repulsion interactions of charged ions have been modelled using interatomic potentials for the corresponding neutral atoms and further details of this can be found in section 2.6 of Chapter 2.

In the present work, we begin by testing the suitability of the FIT(H_N)[102,103,105] and W99[107] force fields (Table 2.1) for modelling chloride salt structures, when supplemented with an appropriate Cl^- potential. Further details on the derivation of the FIT(H_N) and W99 force fields, including the potential parameters for the atoms C, H_c, N, O, H_N, the functional form of the intermolecular potential and the combining rules used can be found in Chapter 2 (section 2.3) of this thesis. In testing the W99 force field, the potential parameters for Cl^- were taken from the fitting of Hejczyk[230], while in the FIT(H_N) force field, two Cl potentials were tested under the assumption that they can be used to approximate the intermolecular interactions involving the Cl^- anion. The first was the Cl potential of Scott & Sheraga[137] (Cl_{SS}) and the second was the potential of Williams[101] (Cl_W) derived from his fitting on organic perchlorohydrocarbon crystal structures. The force fields that result from supplementing these Cl potentials with the FIT(H_N) parameter set will be referred to as FIT(H_N, Cl_{SS}) and FIT(H_N, Cl_W) respectively. Table 6.3 gives the dispersion-repulsion parameters for the Cl_{SS}, Cl_W and Cl^- potentials while Figure 6.14 contrasts the $N^+\cdots Cl^-$ intermolecular potentials derived from the three force fields.

	A (kJ mol^{-1})	B (Å^{-1})	C (kJ mol^{-1} Å^{6})
$Cl_W \cdots Cl_W$	924674.65	3.51	7740.48
$Cl_{SS} \cdots Cl_{SS}$	1313772.53	3.75	10543.65
$Cl^- \cdots Cl$ (HEJCZYK)	1140516.01	3.51	5080.84

Table 6.3: Dispersion-repulsion parameters for the Cl_W, Cl_{SS} and Cl^- (Hejczyk[230]) potentials tested. The parameters for heteroatomic interactions are generated using an arithmetic mean combining law for the B parameter and a geometric mean combining law for the A and C parameters.

A survey of the Cambridge Structural Database (CSD) was performed for organic chloride salts that contain the $NH_3^+Cl^-$ ion pair. The search was performed with V5.30 of the CSD using the November 2008 and February 2009 updates as implemented in ConQuest V1.11. The filters used in the CSD search were: 3D atomic co-ordinates determined, $R \leq 5$ %, not disordered, no errors, no powder structures and only organics. In the search, the $N^+ \cdots Cl$ intermolecular distance was specified as a non-bonded contact of interest and a histogram of the observed $N^+ \cdots Cl$ distances generated (Figure 6.16). The CSD search resulted in a total of 367 chloride crystal structures as well as their associated $N^+ \cdots Cl$ intermolecular distance data. Five crystal structures (Figure 6.15) were chosen from this list of 367 crystal structures for use as a validation set in testing the suitability of the FIT(H_N, Cl_W), FIT(H_N, Cl_{SS}) and W99 force fields. Not all the retrieved chloride structures were suitable and the structures in the validation set were chosen on the basis of three things: all protons involved in hydrogen bonding must have been located from the difference Fourier map, the molecular conformation of the cation must have been rigid and only structures free of solvent molecules were used.

The results following lattice energy minimisation of the training structures as a function of the force field used are given in Table 6.4. All experimental structures were lattice energy minimised using the experimental conformation for the cation (*i.e.* the *ExpMinExp*). Two indicators were used to judge how suitable the force field is for modelling a given test structure. The first was the gross structural reproduction as indicated by the root mean square deviation between the experimental and lattice energy minimised structures following an overlay using a 15-molecule co-ordination sphere ($RMSD_{15}$). The second was the calculated hydrogen bond geometries as indicated by the $N^+ \cdots Cl$ distances of the *ExpMinExp*. These distances were compared with the experimental distances (Table 6.4) of the test structures as well as with the typical CSD distances (Figure 6.16) for chloride salts that contain the $NH_3^+Cl^-$ ion pair.

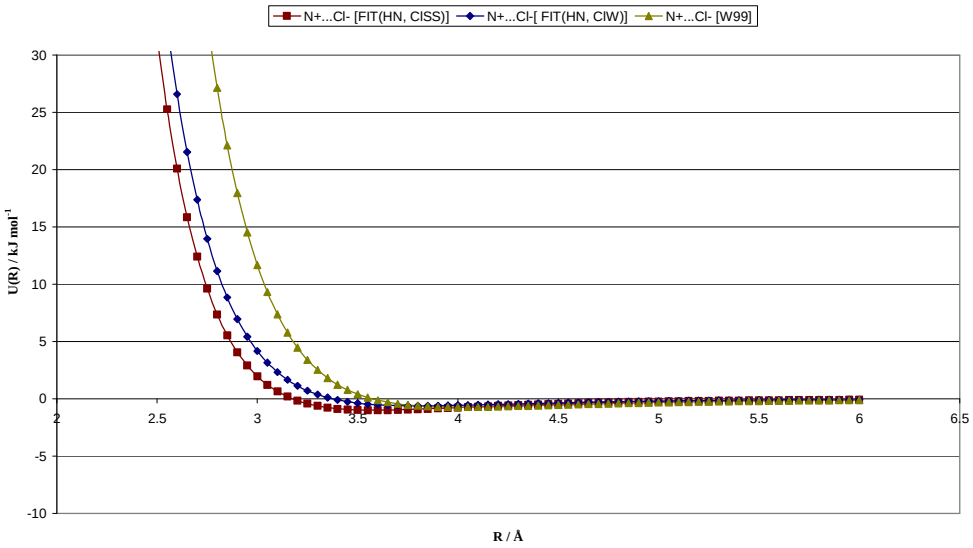


Figure 6.14: Comparison of the $N^+\cdots Cl$ dispersion-repulsion intermolecular potentials derived from the three force fields tested: FIT(H_N, Cl_{SS}), FIT(H_N, Cl_W) and W99. The intermolecular energy was calculated using an *exp*-6 function of the form: $U(R_{ij}) = A_{\iota\kappa} \exp(-B_{\iota\kappa} R_{ij}) - C_{\iota\kappa} / R_{ij}^6$.

If we compare the calculated lattice energy minima of the test set, Table 6.4 shows that the FIT(H_N, Cl_W) force field provides a slightly better description of the gross structural features for all structures when compared to the results obtained using the FIT(H_N, Cl_{SS}) force field. According to the CSD (Figure 6.16), the $N^+\cdots Cl$ distances of organic chloride salts that contain the $NH_3^+Cl^-$ ion pair span the range 3.00-3.30 Å. As can be seen from the calculated $N^+\cdots Cl$ distances given in Table 6.4, both the FIT(H_N, Cl_W) and FIT(H_N, Cl_{SS}) force fields estimate the $N^+\cdots Cl$ distances of **ACEZIR**, **TODCOE** and **ANLINC01** below the minimum value 3.00 Å found in the CSD structures. For those $N^+\cdots Cl$ distances that are calculated within the CSD range of 3.00-3.30 Å, the absolute difference between the experimental and calculated distances was found to be as high as 0.184 Å in **ACEZIR** (-5.77 % decrease) following lattice energy minimisation with the FIT(H_N, Cl_W) force field. With the W99 force field, we find that for structures **TODCOE, ANLINC01** and **CYHACL02**, the force field leads to calculated $N^+\cdots Cl$ distances in the range 3.06-3.18 Å, and this range of calculated values coincides with the CSD range of observed $N^+\cdots Cl$ distances as shown in Figure 6.16. In fact, with the exception of **ACEZIR** and **FAXPOJ,** the W99 force field does a better job in describing the observed hydrogen bond geometries for all training structures. In **ACEZIR** and **FAXPOJ**, significant changes in the positions of

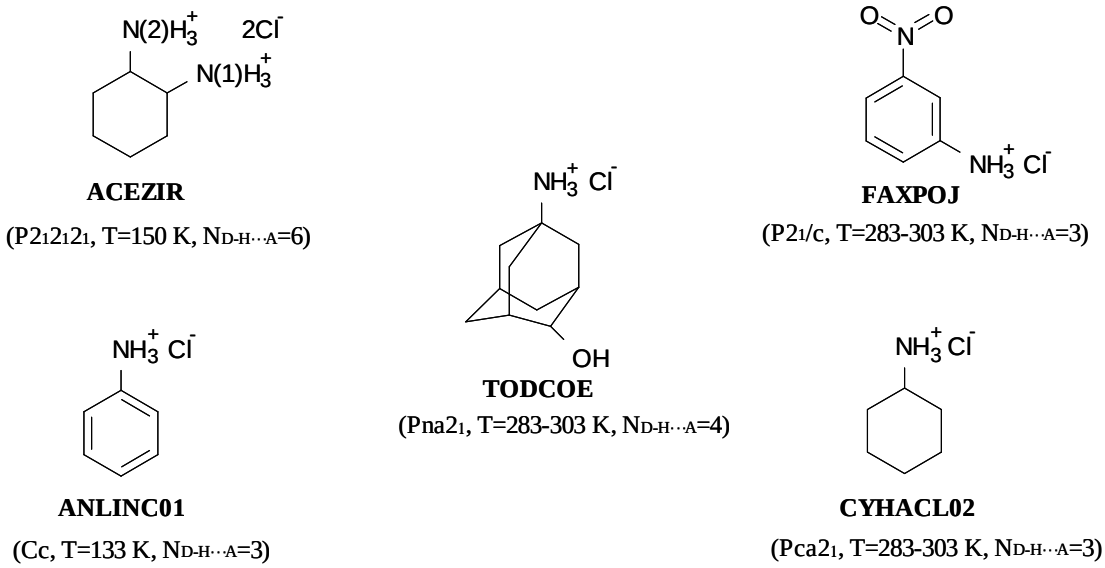


Figure 6.15: Test structures used to validate the FIT(H_N, Cl_W), FIT(H_N, Cl_{SS}) and W99 force fields. The CSD refcodes for the structures are given in bold. $N_{D\text{-}H\cdots A}$ refers to the number of crystallographically distinct hydrogen bonds the cation is involved in. In **TODCOE** there are two N^+-H···Cl, one N^+-H···O and one O-H···Cl hydrogen bonds. The nitrogens of NH_3^+ groups are numbered 1 or 2 in **ACEZIR**. The quoted temperature is the temperature of the crystal structure determination.

the chloride ions were observed following lattice energy minimisation with the W99 force field and this explains the large errors in the gross structural reproduction as well as the calculated hydrogen bond geometries for the two structures.

The conclusion that can be drawn from the potential validation on this limited set of structures is that when the results are viewed holistically, no single force field adequately describes the intermolecular interactions found in all the organic chlorides structures considered. The W99 force field reproduces the test structures with realistic hydrogen bond geometries and gross packing features for three of the five structures considered. The key problem with the FIT(H_N, Cl_W) and FIT(H_N, Cl_{SS}) force fields is that they both underestimate the $N^+\cdots Cl$ distances in the test structures below the CSD range of 3.00-3.30 Å. These trends are consistent with Figure 6.14, which shows the variation in the $N^+\cdots Cl$ intermolecular energy with the non-bonded distance *R* as a function of the force field used. In the case of both the FIT(H_N, Cl_{SS}) and FIT(H_N, Cl_W) force fields, the shallow repulsive wall of the $N^+\cdots Cl$ intermolecular potential means that the intermolecular energy increases by only 5.40 kJ mol^{-1} and 6.98 kJ mol^{-1} respectively when *R* is changed from 3.00 Å to 2.98 Å. With the W99 force

Structure	Intermolecular potential	Intermolecular $N^+\cdots Cl$ distances (Å)		a (Å) [% Δa]	b (Å) [% Δb]	c(Å) [% Δc]	β (°) [% Δβ]	ρ (g cm^{-3}) [% Δρ]	U (kJ mol^{-1})	F	$RMSD_{15}$ (Å)
		N1	N2								
ACEZIR	-	3.187, 3.147, 3.092	3.185, 3.130, 3.113	8.1138(2)	9.5567(2)	11.9553(3)	90	1.341	-	-	-
ExpMinExp	DMA+ FIT(H_N, Cl_W)	3.003, 3.016, 2.968	3.060, 3.052, 2.940	0.48	-3.84	-3.68	-	7.46	-869.91	41.77	0.220
	DMA+ FIT(H_N, Cl_{SS})	2.940, 2.968, 2.919	3.001, 2.998, 2.880	-0.20	-5.76	-5.15	-	12.09	-899.58	86.44	0.316
	DMA+ W99	3.394, 3.403, 3.331	3.340, 3.299, 3.315	-0.67	4.35	4.14	-	-7.35	-793.85	125.29	0.527
FAXPOJ		3.229, 3.169, 3.116	-	9.476(1)	9.732(1)	8.459(1)	78.76(1)	1.516	-	-	-
ExpMinExp	DMA+ FIT(H_N, Cl_W)	3.045, 3.052, 3.013	-	-0.60	-5.42	-0.73	-0.17	7.20	-754.94	39.56	0.221
	DMA+ FIT(H_N, Cl_{SS})	3.001, 3.007, 2.968	-	-1.45	-7.14	-1.59	-0.59	11.23	-785.77	70.52	0.291
	DMA+ W99	3.362, 3.330, 3.360	-	1.58	1.61	1.17	-0.39	-4.12	-694.46	13.17	0.194
TODCOE		3.214, 3.147	-	14.3910(10)	10.2370(10)	7.0560(10)	90	1.301	-	-	-
ExpMinExp	DMA+ FIT(H_N, Cl_W)	3.010, 2.979	-	0.50	-1.04	-4.37	-	5.15	-728.46	26.67	0.188
	DMA+ FIT(H_N, Cl_{SS})	2.958, 2.937	-	-0.11	-1.80	-5.80	-	8.22	-757.36	46.26	0.249
	DMA+ W99	3.145, 3.099	-	1.84	-0.37	-4.05	-	2.71	-676.22	25.56	0.182
ANLINC01		3.145, 3.087, 3.102	-	15.735(3)	5.3094(12)	8.397(2)	101.029(6)	1.250	-	-	-
ExpMinExp	DMA+ FIT(H_N, Cl_W)	2.998, 2.972, 2.974	-	-4.66	-6.07	0.81	-1.73	10.16	-727.99	71.68	0.224
	DMA+ FIT(H_N, Cl_{SS})	2.955, 2.926, 2.929	-	-6.79	-6.87	-0.49	-1.78	15.11	-756.05	108.95	0.257
	DMA+ W99	3.087, 3.065, 3.064	-	-2.21	-4.66	3.72	-1.52	2.90	-680.52	53.12	0.208
CYHACL02		3.219, 3.162, 3.187	-	9.339(2)	11.449(2)	7.546(1)	90	1.117	-	-	-
ExpMinExp	DMA+ FIT(H_N, Cl_W)	3.044, 3.027, 3.058	-	-0.12	-1.74	-5.44	-	7.75	-694.73	40.31	0.186
	DMA+ FIT(H_N, Cl_{SS})	2.984, 2.974, 2.973	-	-4.10	-2.54	-4.34	-	11.84	-721.43	52.71	0.249
	DMA+ W99	3.179, 3.140, 3.168	-	-0.30	-5.63	0.44	-	5.81	-651.32	56.15	0.271

Table 6.4: Results following lattice energy minimisation of the validation structures as a function of the force field used. For the calculated lattice energy minima, the % change in the calculated cell density and lattice parameters are given below the experimental values. DMA refers to the distributed multipole model of the *ab initio* charge distribution for the cation used in all modelling.

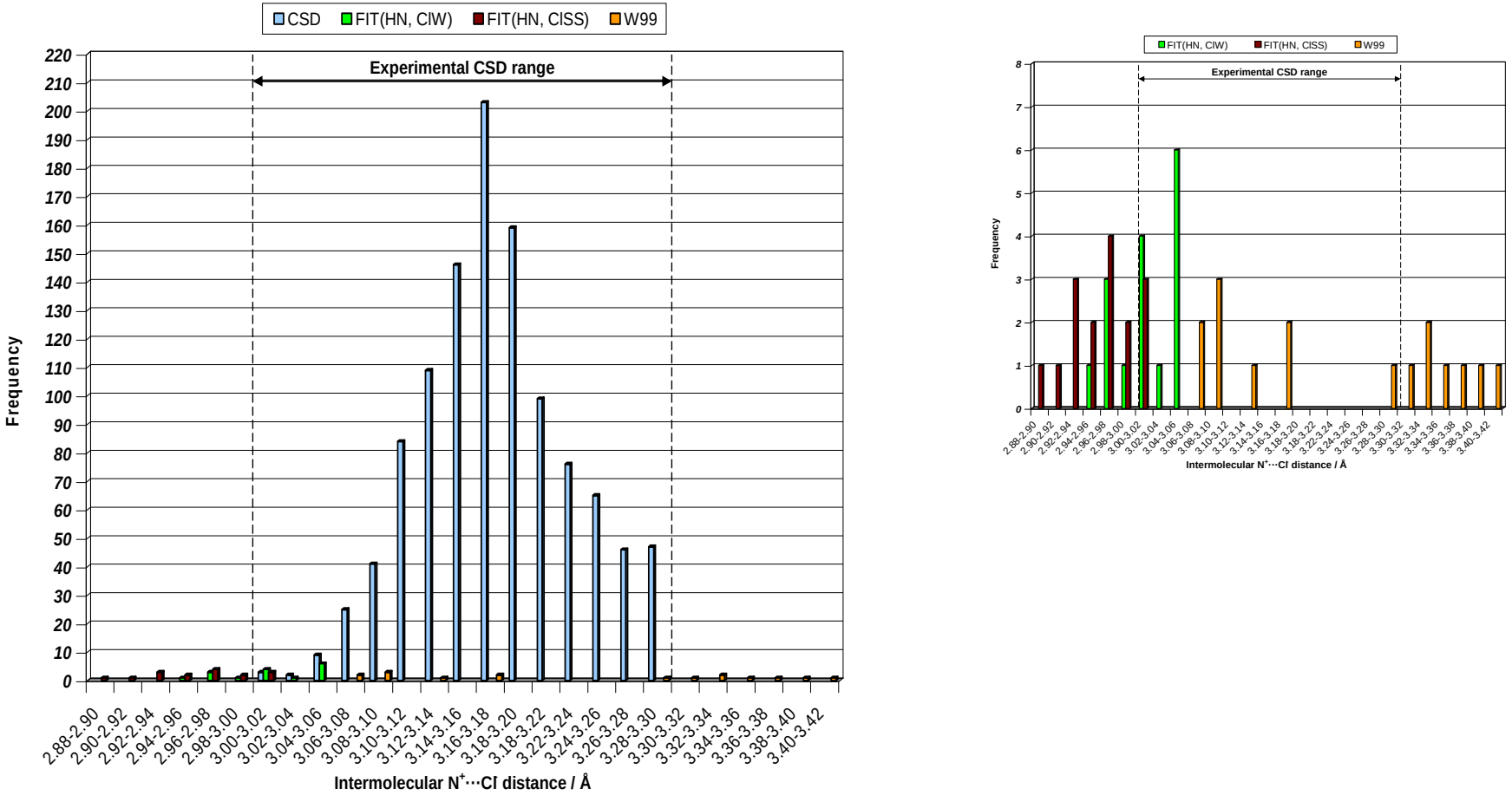


Figure 6.16: Comparison of the observed $N^+\cdots Cl$ intermolecular distances found in $NH_3^+Cl^-$ containing salt structures of the CSD and those calculated for the five test structures following use of the FIT(H_N, Cl_W), FIT(H_N, Cl_{SS}) or W99 force fields in conjunction with an explicit distributed multipole electrostatic model. The experimental CSD range for the intermolecular $N^+\cdots Cl$ distance is illustrated as lying within the range 3.00-3.30 Å. An expanded version of the histogram emphasising the relative frequency of the $N^+\cdots Cl$ data points from the FIT(H_N, Cl_W), FIT(H_N, Cl_{SS}) and W99 minimisations is shown to the right.

Amantadine hydrochloride	Exp (FINVAZ)	ExpMinExp	% change	ExpMinOpt	% change
a (Å)	20.549	21.030	2.34	20.549	0.54
b (Å)	11.138	11.140	0.02	11.077	-0.55
c (Å)	9.658	9.633	-0.26	9.823	1.71
α (°)	90	90	-	90	-
β (°)	108.81	112.33	3.24	109.60	-
γ (°)	90	90	-	90	-
Volume (Å^3)	2092.418	2087.514	-0.23	2117.812	1.21
r (N$^+$···Cl) / Å	3.172, 3.163, 3.219	3.144, 3.154, 3.139	-	3.134, 3.158, 3.181	-
E_{latt} (kJ mol^{-1})	-	-625.75	-	-626.05	-
F	-	25.69	-	5.85	-
$RMSD_{15}$ / Å	-	0.203	-	0.080	-
Amantadine hydrochloride	**Exp (aman_103_0m)**	**ExpMinExp**	**% change**	**ExpMinOpt**	**% change**
a (Å)	23.091	23.144	0.23	23.249	0.69
b (Å)	11.170	11.008	-1.45	11.077	-0.83
c (Å)	25.320	25.644	1.28	25.681	1.42
α (°)	90.00	90.00	-	90.00	-
β (°)	107.46	106.56	-0.84	106.12	-
γ (°)	90.00	90.00	-	90.00	-
Volume (Å^3)	6229.500	2087.387	0.52	2117.823	-1.25
r (N$^+$···Cl) / Å	3.163, 3.179, 3.198, 3.202, 3.175, 3.157, 3.181, 3.211, 3.175	3.121, 3.164, 3.151, 3.156, 3.139, 3.128, 3.138, 3.170, 3.159	-	3.134, 3.181, 3.158, 3.181, 3.134, 3.158, 3.134, 3.181, 3.158	-
E_{latt} (kJ mol^{-1})*	-	-630.51	-	-626.05	-
F	-	15.22	-	40.58	-
$RMSD_{15}$ / Å	-		-		-
Memantine hydrochloride	**Exp (tinr3c0m)**	**ExpMinExp**	**% change**	**ExpMinOpt**	**% change**
a (Å)	28.164	27.746	-1.48	27.6409	-1.86
b (Å)	28.164	27.746	-1.48	27.6409	-1.86
c (Å)	8.420	8.515	1.13	8.695	3.27
α (°)	90	90	-	90	-
β (°)	90	90	-	90	-
γ (°)	120	120	-	120	-
Volume (Å^3)	5783.643	5566.007	-1.85	5643.623	-0.53
r (N$^+$···Cl) / Å	3.148, 3.127, 3.132	3.184, 3.096, 3.103	-	3.172, 3.115, 3.114	-
E_{latt} (kJ mol^{-1})	-	-615.42	-	-611.70	-
F	-	7.14	-	20.94	-
$RMSD_{15}$ / Å	-	0.112	-	0.155	-

Table 6.5: The calculated *ExpMinExp* and *ExpMinOpt* structures for anhydrous amantadine hydrochloride (**FINVAZ**, *aman_103_0m*) and memantine hydrochloride (*tinr3c0m*). * The lattice energy is quoted per mole of the ions.

field, the same change in the $N^{+}\cdots Cl^{-}$ non-bonded distance leads to a 15.47 kJ mol^{-1} increase in the intermolecular energy. These differences in the $N^{+}\cdots Cl^{-}$ intermolecular energies explain why it is easier to populate intermolecular distances below the CSD minimum of 3.00 Å with the FIT(H_N, Cl_{SS}) and FIT(H_N, Cl_W) force fields when compared to the results obtained with the W99 force field.

Given that the W99 force field was the most promising of the three force fields tested, it was used to calculate the lattice energy minima of the anhydrous crystal structures for amantadine (**FINVAZ** and *aman_103_0m*) and memantine (*tinr3c0m***)** hydrochloride using the experimental (*ExpMinExp*) and *ab initio* optimised (*ExpMinOpt*) conformations for the cation. Table 6.5 shows a comparison of the experimental structures of amantadine hydrochloride, **IX**, and memantine hydrochloride, **X**, with the calculated *ExpMinExp* and *ExpMinOpt* structures for each salt. In both **IX** and **X**, the % change in the cell parameters is within 4 % of the experimental value for the *ExpMinExp* and *ExpMinOpt* structures. All the calculated $N^{+}\cdots Cl^{-}$ distances are within the CSD range of 3.00-3.30 Å. Both the F and $RMSD_{15}$ values of the *ExpMinExp* and *ExpMinOpt* are relatively low. Given the above successes in using the W99 force field to model the anhydrous crystal structures of **IX** and **X** and in light of the success achieved with three of the five CSD structures tested, the force field was used in the crystal structure prediction of amantadine hydrochloride and memantine hydrochloride.

7 Blind crystal structure prediction of 1,8-naphthyridinium fumarate

7.1 Introduction

So far in this thesis, we have seen examples of where the modelling has predicted the experimental carboxylate (Chapter 5, section 5.4.1) or chloride (Chapter 6, section 6.3.2) salt structure as the most thermodynamically stable in the crystal energy landscape, in agreement with the results of experimental screens which have indicated the experimental structure as the only solid form or the most stable polymorph. Whilst these successes have proven useful in validating that the computational model is capable of predicting organic salt crystal structures, the prior knowledge of the experimental structures has meant that such studies cannot be regarded as blind predictions. The aim of this chapter is to put the model to the test by participating in the fifth international blind test of crystal structure prediction and performing the calculations under truly blind conditions. The candidate system for the blind prediction will be the 1,8-naphthyridinium fumarate salt (Figure 7.1), **XIX**. In numbering this solid form, we have used the conventions of the international blind tests (19th system proposed to date) rather than the running order of structures in this thesis. The work of Chapters 4 and 5 give us the necessary foundation to take on this blind test challenge. This is because of the close similarity in the molecular structures of the ions in **XIX** and those of the pyridinium carboxylate salts studied in these chapters.

Since 1999, the Cambridge Crystallographic Data Centre (CCDC) has organised the international blind tests in crystal structure prediction. In these blind tests, all research groups actively developing a computational method of organic crystal structure prediction are sent the molecular diagrams of candidate systems, and more recently, the

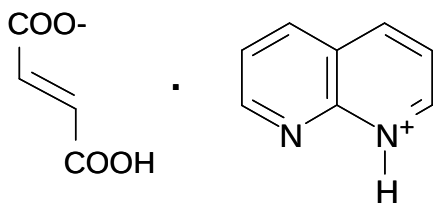


Figure 7.1: The molecular structure of the 1,8-naphthyridinium fumarate salt, **XIX**.

relevant crystallisation conditions. Participants are asked to submit three predictions for the likely crystal structure before a given deadline. The blind tests, which are held every 2-3 years, allow an unbiased assessment of the progress made in methods of *ab initio* crystal structure prediction. Similar blind test have been used to assess progress in other areas of predictive modelling and examples of these include the areas of protein folding[244] and solvation energies[245]. The fifth international blind test of crystal structure prediction contains six candidate systems, each with its own particular challenges for the computational modelling. The type of molecular structures posed as candidates for the blind tests reflect perceived areas of difficulty in modelling organic crystal structures (*i.e.* molecular flexibility, $Z' > 1$) as well as the current interests of crystal engineers. This last point is illustrated by the fact that it was only in the fourth blind test of crystal structure prediction where the participants were challenged with predicting the experimental structure of a cocrystal and this reflects the rapid growth in the interest of cocrystal solid forms since the first[246] international blind test. For the first four[212,239,246,247] blind tests, the target systems chosen belonged to one of the following categories. Category four - of which **XIX** is a member - was introduced in the fourth blind test while categories one to three were used in the first three blind tests.

1) *Small, rigid molecules*: Only the elements C, H, N and O, < 25 atoms
2) *Rigid molecules*: Must contain elements/functional groups that present a challenge for modelling methods, *ca* 30-40 atoms.
3) *Molecules with several degrees of conformational flexibility*: Usually rotation is about exocyclic bonds.
4) *Two component crystal of rigid molecules*: This includes salts, cocrystals and solvates.

The most recent and fifth international blind test of crystal structure prediction was announced in November 2009 and all predictions had to be submitted by August 2010. The blind test was organised with the intention of building upon the successes of the previous blind tests and as such includes two new categories in addition to the ones listed above. These are:

5) *Flexible molecule with 4-8 internal degrees of freedom*: Z'≤2, any space group, 50-60 atoms.
6) *Molecule where more than one polymorph is known*: Should roughly fall into one of the first four categories.

In all the previous blind tests, the most widely used hypothesis by participants was that the observed crystal structure would be the most thermodynamically stable in lattice energy. This hypothesis is not always justified however, and according to Ostwald's Law of Stages[38], the first crystallisation from the melt or solution will yield a metastable crystalline form. In the present work, we will continue to use the lattice energy to rank the predicted structures and assume that the crystallisation is under thermodynamic control.

7.2 Method

The molecular structures of the 1,8-naphthyridinium and fumarate ions were extracted from the crystal structures of **NETLUT**[248] and **RABYID**[249] respectively and used in separate gas phase geometry optimisations at the MP2/6-31G(d,p) level of theory using the GAUSSIAN03[119] suite of programs. For the 1,8-naphthyridinium ion, there are no flexible torsional degrees of freedom and so the ion was optimised freely in the gas phase. The fumarate anion displays torsional flexibility about the COOH, COO^- and OH groups (Figure 7.2). Initially, we chose to do a relaxed one-dimensional conformational scan about each of these torsions in steps of 20 °. This was followed by a two-dimensional conformational scan (in steps of 20 °) about the torsions defining rotation about the OH and COOH groups and the resulting potential energy surface led to the identification of four distinct conformational minima [denoted s-*cis*(1), s-*cis*(2), s-*trans*(1) and s-*trans*(2)] of the fumarate ion. Each of these fumarate conformations were used in separate rigid body searches for hypothetical crystal structures of **XIX**.

CrystalPredictor[88] was used to generate hypothetical crystal structures of **XIX**. For each conformation of the fumarate ion, a search for stable packings was performed in the space groups with highest occurrence frequency[250] for organic molecules: $P1$, $P\bar{1}$, $P2_1$, $P2_1/c$, $P2_12_12$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$, $Pbca$, $Pbcn$, Cc, $C2$, $C2/c$, $C2/m$, $P2_1/m$. A total of 100,000 trial crystal structures were generated in the above

space groups for each of the fumarate conformations. This was followed by another search in the remaining high symmetry space groups within the current capabilities of CrystalPredictor: Cm, Pc, $P2/c$, $C222_1$, $Pmn2_1$, $Cmc2_1$, $Aba2$, $Fdd2$, $Iba2$, $Pnna$, $Pccn$, $Pbcm$, $Pnnm$, $Pmmn$, $Pnma$, $Cmcm$, $Cmca$, $Fddd$, $Ibam$, $P4_1$, $P4_3$, $I\bar{4}$, $P4/n$, $P4_2/n$, $I4/m$, $I4_1/a$, $P4_12_12$, $P4_32_12$, $P\bar{4}2_1c$, $I\bar{4}2d$, $P3_1$, $P3_2$, $R3$, $P\bar{3}$, $R\bar{3}$, $P3_121$, $P3_221$, $R3c$, $R\bar{3}c$, $P6_1$, $P6_3$, $P6_3/m$, $P2_13$ and $Pa\bar{3}$. For each fumarate conformation, the search in the above high symmetry space groups also generated a total of 100,000 trial crystal structures. As before, convergence was assessed on the basis of the number of times each of the low energy structures was found by CrystalPredictor. For the CrystalPredictor searches involving the s-*cis*(1) and s-*trans*(1) fumarate conformations, the global minimum structures were found twenty-seven and ten times respectively under simulations that were limited to the statistically most common space groups. For the searches involving the s-*cis*(2) and s-*trans*(2) fumarate conformations, the global minimum structures were found forty-six and six times respectively when the searches were performed in the statistically most common space groups. The variations in the search redundancies suggest that more CrystalPredictor minimisations would have been beneficial for some fumarate conformations [most notably s-*trans*(1) and s-*trans*(2)] but this was limited by the available computational resources at the time. Following an initial validation of the FIT(H_O,H_N) force field (Chapter 2, section 2.3) on **RABYID**, the force field was found to be adequate and used in the search for hypothetical crystal structures of **XIX**. In the CrystalPredictor searches, the electrostatic interactions were modelled using the atomic charges derived from fitting (CHELPG scheme[93]) to the electrostatic potential of the *ab initio* wavefunctions for the ions. The 1000 lowest energy structures produced from each search were passed to DMACRYS[89] for rigid body lattice energy minimisations using a distributed multipole model of the MP2/6-31G(d,p) *ab initio* wavefunction for the isolated ions. The program GDMA2.1[195] was used to perform the distributed multipole analysis of the *ab initio* wavefunctions. Only structures at true lattice energy minima were kept and these structures were subsequently clustered to remove multiple findings of the same lattice energy minima. The clustering was done by comparing the relative lattice energies and cell densities, simulated X-ray powder diffraction patterns and COMPACK overlays of the 15 molecule co-ordination spheres ($RMSD_{15}$) using a

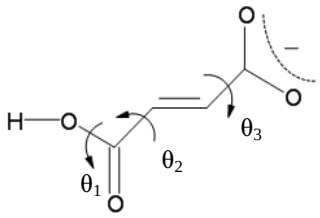


Figure 7.2: Illustration of the flexible torsion angles on the fumarate ion.

similarity threshold of 0.1 Å or less for identical structures.

Following an estimate of the rigid body lattice energies for the hypothetical structures using a distributed multipole model, the most stable structures within 12 kJ mol^{-1} of the global minimum structure were re-minimised with CrystalOptimizer[116] (Chapter 2, section 2.5.3). The program accounts for molecular flexibility by minimising the lattice energy ($E_{latt}=U_{inter} + \Delta E_{intra}$) with respect to the cell parameters, molecular positions and orientations as well as selected torsion angles. For the 1,8-naphthyridinium fumarate salt, the torsion angles optimised were those defining rotation about the OH (θ_1), COOH (θ_2) and COO^- (θ_3) groups (Figure 7.2). The intermolecular energy was computed as in the rigid body model described above but the MP2/6-31G(d,p) atomic multipole moments were rotated with the local environment and re-calculated if the change in a given torsion angle was ≥ 4 °. The intramolecular energy was calculated using a Local Approximate Model (LAM) based on a second order Taylor expansion and the LAM was updated with the distributed multipole model. The reference point for the Taylor expansion was the single point intramolecular energy for an ion calculated at the MP2/6-31G(d,p) level of theory. Both the rotated multipoles and intramolecular conformations from the LAM were stored in a database and used in subsequent refinements involving other hypothetical crystal structures.

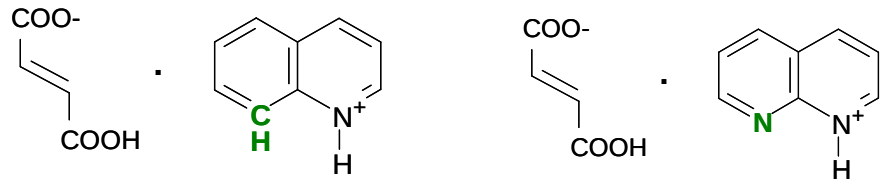
All packing diagrams, hydrogen bond motifs and associated graph sets were calculated using Mercury CSD 2.3[218]. Surveys of the Cambridge Structural Database (CSD) were performed with V5.30 of the CSD using all updates up to September 2009. The filters used in the CSD searches were: 3D co-ordinates determined, not disordered, no errors, no powder X-ray diffraction structures and only organics. For structures with more than one entry in the database, the best determination (as measured by the

refinement residual, R_1) was used to extract the numerical values for selected torsion angles of the fumarate ion(s).

7.3 Results and discussion

7.3.1 CSD survey of structures related to the 1,8-naphthyridinium fumarate salt

A survey of the Cambridge Structural Database (CSD) was performed to identify the number of salt structures that contain either the monoionised fumarate ion or the 1,8-naphthyridinium ion. The search fragments used in the separate CSD searches were the molecular structures of the ions shown in Figure 7.1. When searching for salts that contain the fumarate ion, it was necessary to define the text string "fumarate" along with the molecular structure of the ion so as to eliminate the possibility of retrieving structures with the isomeric maleate ion. A total of 28 structures were found following the search for salts that contain the monoionised fumarate ion and the CSD refcodes for these structures can be found in Table 7.4 of the appendix (see section 7.7.2). Some of these fumarate salts had one or more neutral fumaric acid molecules in the crystal lattice. We have already seen examples of such structures in the 2:1:1 and 2:1:2 salts of 4-dimethylaminopyridinium fumarate-fumaric acid reported in Chapter 4 (see section 4.3.7). A similar search for salts containing the 1,8-naphthyridinium ion returned only one hit and this was a structure (CSD refcode: **NETLUT**) with an inorganic tetraphenylborate counterion.



RABYID **XIX**

Figure 7.3: Molecular structures of the quinolinium fumarate salt, **RABYID** (left), and the 1,8-naphthyridinium fumarate salt, **XIX** (right). Shown in green are the key distinguishing atoms in the molecular structures of **RABYID** and **XIX**.

Among the 28 fumarate salts retrieved from the CSD was the crystal structure of quinolinium fumarate, **RABYID** (Figure 7.3). Of all the CSD entries surveyed, the molecular structures of the ions in **RABYID** were the most closely related to those of the 1,8-naphthyridinium fumarate salt, **XIX**. The key distinguishing features in the molecular structures of the two salts (Figure 7.3) is that **RABYID** has a C-H group in place of the neutral N atom in the 1,8-naphthyridinium ion of **XIX**. The crystal packing of **RABYID** (Figure 7.4) is based on ribbons [graph set: $C_1^1(7)$] of fumarate ions and these ions interact with the quinolinium cation via an $R_2^2(7)$ dimer motif. The crystal structure of **RABYID** serves as a good candidate for validating the force field used. We already know from the work of Chapter 4 (section 4.4.1) that the FIT(H_O, H_N) force field is a suitable starting point for modelling the crystal structures of pyridinium carboxylate salts. However, it is worth investigating how the force field performs for carboxylate salts with larger cations such as the quinolinium cation.

The crystal structure of **RABYID** was taken from the CSD and used to calculate the *ExpMinExp* and *ExpMinOpt* structures. The estimated lattice energies and cell parameters of the calculated *ExpMinExp* and *ExpMinOpt* are given in Table 7.1. The table shows that using the FIT(H_O,H_N) force field in combination with a distributed multipole electrostatic model leads to calculated lattice energy minima with cell parameters that are within 3 % of the experimental values for the *ExpMinExp* and approximately 5 % for the *ExpMinOpt*. Similar errors were observed in the crystallographic parameters of salts reported in Chapter 4. The $N^+\cdots O^-$ intermolecular distances in the calculated lattice energy minima were reproduced within 5 % of the experimental values. In light of the above results for **RABYID**, the FIT(H_O,H_N) force field was judged as suitable for use in the blind crystal structure prediction of 1,8-naphthyridinium fumarate.

7.3.2 Conformational scans of the flexible torsion angles in fumarate ion

A relaxed conformational scan was performed of the torsion defining rotation around the COOH group, θ_2. Three conformational scans were performed according to one of the following levels of theory: HF, MP2 and DFT (PBE functional), using the 6-31G(d,p) basis set. The different levels of theory were used to provide re-assurance of the estimated relative energies in the conformational minima. We note however, that

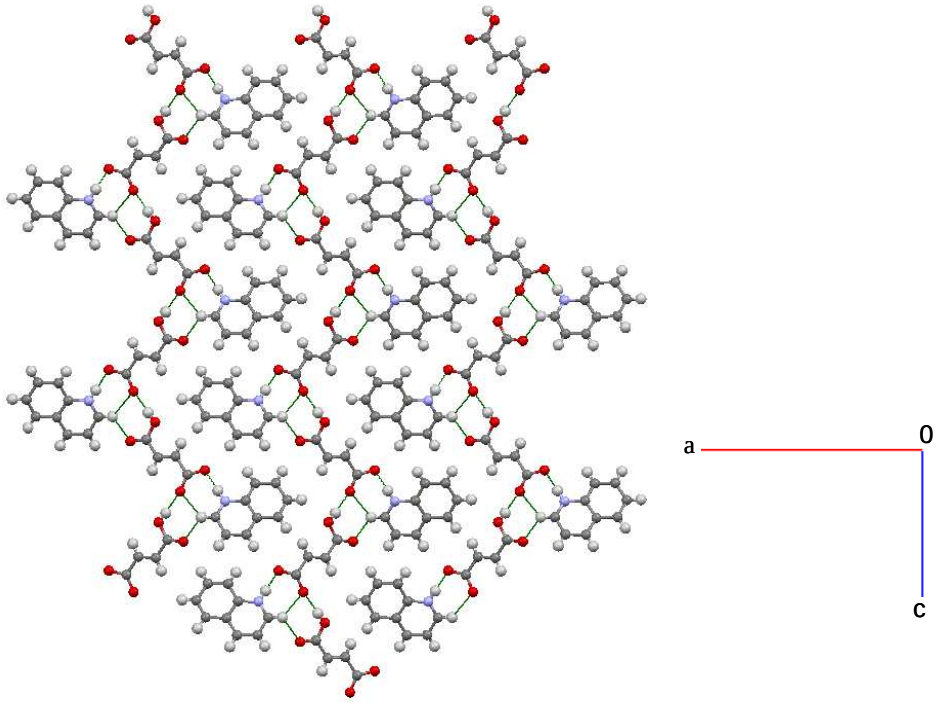


Figure 7.4: Crystal packing found in **RABYID** as viewed down the *b*-axis.

	RABYID	**ExpMinExp**	**% change**	**ExpMinOpt**	**% change**
a (Å)	22.5838(5)	22.27	-1.41	21.71	-3.86
b (Å)	3.72730(10)	3.82	2.42	3.80	2.08
c (Å)	13.2912(5)	13.46	1.27	13.97	5.10
α (°)	90	90	-	90	-
β (°)	90	90	-	90	-
γ (°)	90	90	-	90	-
Volume (Å^3)	1118.81(6)	1143.98	2.25	1154.08	3.15
O···O / Å	2.667	2.748	3.03	2.790	4.61
N$^+$···O / Å	2.553	2.666	4.43	2.659	4.15
E_{latt} (kJ mol^{-1})	-	-616.29	-	-588.12	-
F	-	22.05	-	90.41	-
$RMSD_{15}$ / Å	-	0.24	-	0.33	-

Table 7.1: Lattice energy and crystallographic parameters for the calculated *ExpMinExp* and *ExpMinOpt* structures of **RABYID**.

since the crystal structure prediction of **XIX** was performed using gas phase optimised conformations and multipole moments derived from the MP2/6-31G(d,p) wavefunctions for the ions, only the relative energies from the MP2/6-31G(d,p) conformational scans will be used to estimate the lattice energy in the equation $E_{latt}=U_{inter} + \Delta E_{intra}$.

The results from the conformational scan of θ_2 are shown in Figure 7.5. The figure shows that there are two distinct conformational minima for the fumarate ion. The first is the s-*trans* conformation [θ_2 = -180 °] and this is the conformation adopted by the fumarate ion of **RABYID**. The second conformational minimum occurs at a torsion angle of 0 ° and this minimum is identified as the s-*cis* conformation of the fumarate ion. The notation of s-*cis* or s-*trans* refers to the spatial arrangement of two conjugated double bonds as defined in the Gold Book of the International Union of Pure and Applied Chemistry. The relative energies of the s-*cis* and s-*trans* conformations vary according to the level of theory used but in all cases the s-*cis* conformation is more stable than the s-*trans*. The survey of conformations adopted by the fumarate ions of salt structures retrieved from the CSD (Table 7.4) has shown that 21 of the 32 fumarate ions surveyed had the s-*cis* conformation while 11 had the s-*trans* conformation. The relative energies of the two conformations from the HF (0.43 kJ mol^{-1}) and DFT-PBE (0.47 kJ mol^{-1}) methods are comparable but the estimate from the MP2 (0.93 kJ mol^{-1}) is roughly twice as large as either of these. Despite this fact, visual inspection of Figure 7.5 shows that there is good overlap of the potential energy curves in the vicinity of the minima but the overlap is poorer around the energy maxima.

A similar conformational scan of the COO^- (θ_3) torsion has shown that rotation by 180 ° leads to a conformational minimum with the same energy as the starting planar conformation at scan angle of 0 °. This is what we would expect given the electron delocalisation of the charge in the carboxylate group, which leads to approximately identical electronic environments on the two oxygen atoms.

The only remaining flexible torsion we need to consider is θ_1 which defines rotation around the OH acidic proton of the fumarate ion. Figure 7.6 shows a contour plot of the results following a two dimensional conformational scan [MP2/6-31G(d,p) level of theory] of the torsions θ_1 and θ_2. The relative energies of the different conformations are distinguished by colour and the white circles correspond to conformations that are observed in experimental fumarate ions retrieved from the CSD (Table 7.4). The global minimum conformation is found at the co-ordinates

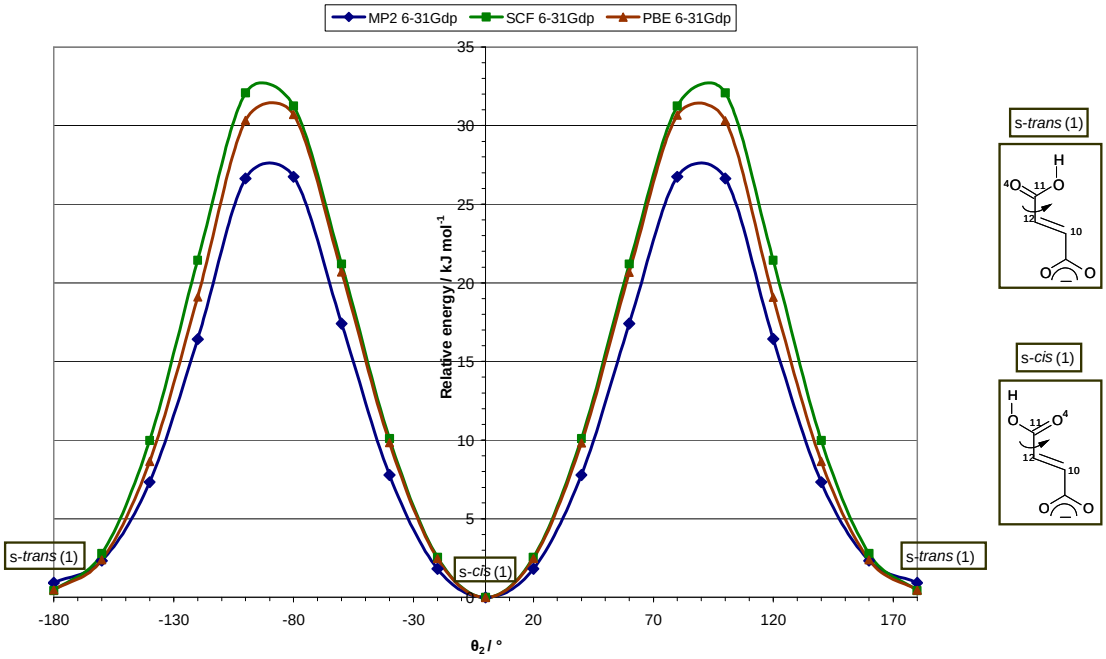


Figure 7.5: Potential energy curve for rotation about the C10-C12-C11-O4 carboxylic acid torsion of the fumarate ion, θ_2.

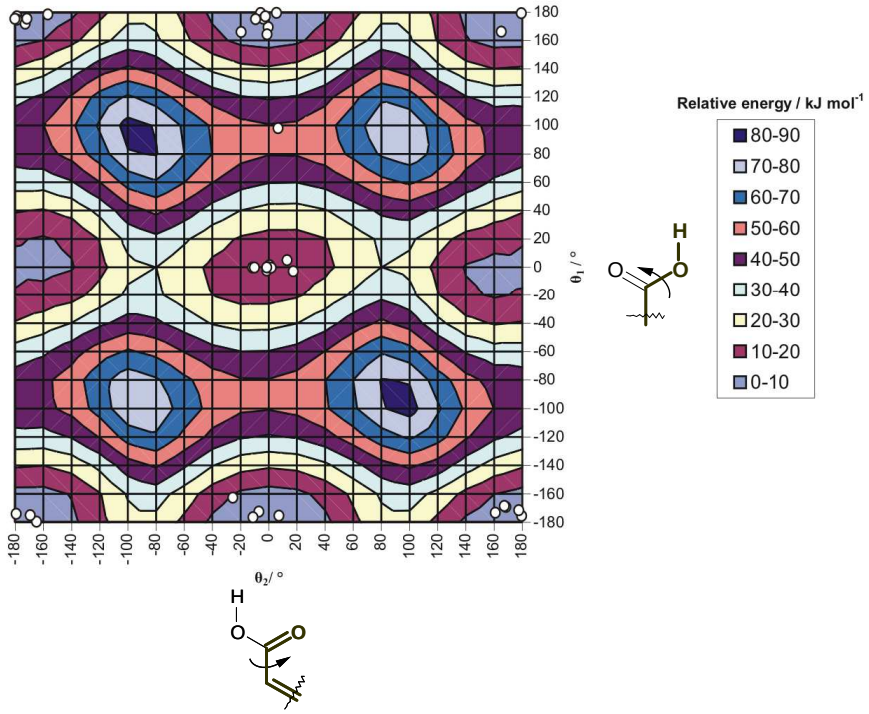
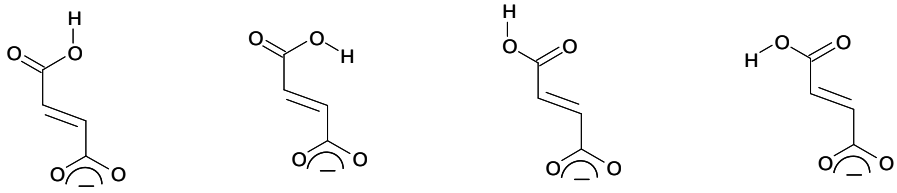


Figure 7.6: Illustration of the contour surface that results following a relaxed two dimensional conformational scan of the torsions defining rotation around the OH, θ_1, and COOH, θ_2, groups. All conformational energies are relative to the energy of the global minimum (θ_1=-180 °, θ_2=0 °) conformation. White circles correspond to observed conformations in fumarate ions retrieved from the CSD. The conformational scan was performed at the MP2/6-31G(d,p) level of theory.



	s-*trans*(1)	s-*trans*(2)	s-*cis*(1)	s-*cis*(2)
(θ_1,θ_2):	(-180°,-180°)	(-5.6°,-158.8°)	(-180°,0°)	(0°,0°)
# in CSP*:	59	0	64	30
ΔE_{intra} / kJ mol^{-1}:	0.93	6.65	0	13.08
Lowest E_{latt} / kJ mol^{-1}:	-601.24	-591.31	-600.67	-604.44

Figure 7.7: Illustration of the four conformational minima of the fumarate ion used in the rigid body searches for hypothetical crystal structures. The ΔE_{intra} is the intramolecular energy of a particular conformation relative to the intramolecular energy of the most stable s-*cis*(1) conformation, calculated at the MP2/6-31G(d,p) level of theory. * Refers to the number of predicted structures in the rigid body crystal energy landscape (Figure 7.8) of **XIX** with that particular conformation.

(-180°, 0°) in (θ_1, θ_2) and corresponds to the s-*cis* conformation of the fumarate ion. The second most stable conformation with a relative energy of 0.93 kJ mol^{-1} can be found at the co-ordinates (-180°, -180°) and corresponds to the s-*trans* conformation. Both these conformational minima have the OH acidic proton pointing away from the fumarate ion and are denoted s-*cis*(1) and s-*trans*(1) in Figure 7.7. The two remaining conformational minima occur at the co-ordinates (0°, 0°) and (-5.6°, -158.8°) in (θ_1, θ_2) and have relative energies of 13.08 kJ mol^{-1} and 6.65 kJ mol^{-1} with respect to the global minimum s-*cis*(1) conformation. These conformations are denoted s-*cis*(2) and s-*trans*(2) respectively in Figure 7.7. Of the four conformational minima illustrated in Figure 7.7, s-*trans*(2) is the only one that is non-planar and the only one that is not observed among the 28 fumarate salts retrieved from the CSD.

7.3.3 Rigid body crystal energy landscape of 1,8-naphthyridinium fumarate

The combined results of the rigid body CrystalPredictor searches for hypothetical crystal structures of 1,8-naphthyridinium fumarate, **XIX**, are shown in Figure 7.8. In the figure, the lattice energy minima are denoted according to the fumarate conformation and the graph set of the hydrogen bond motif. Structures with the

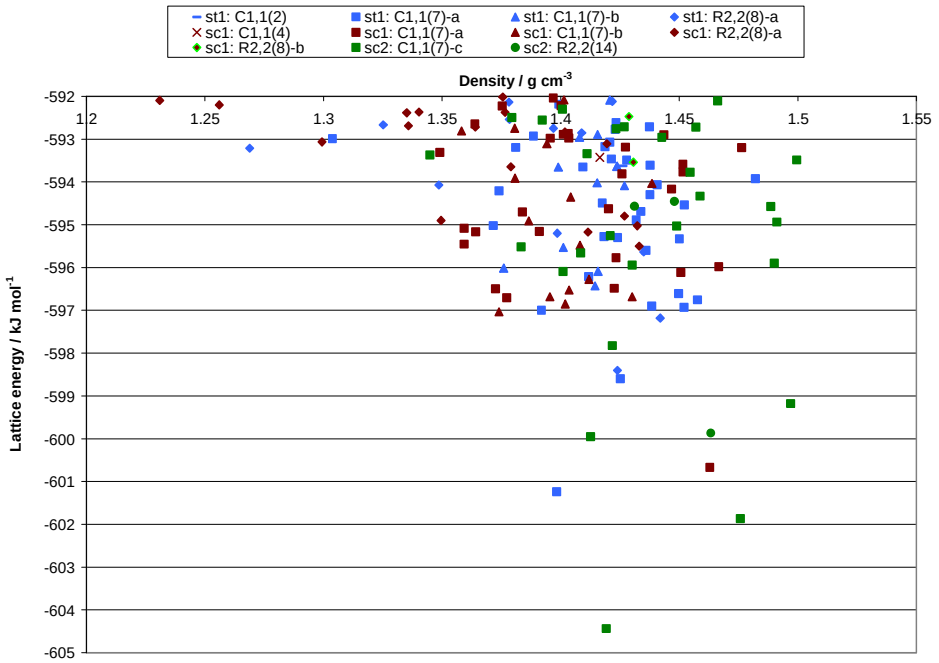


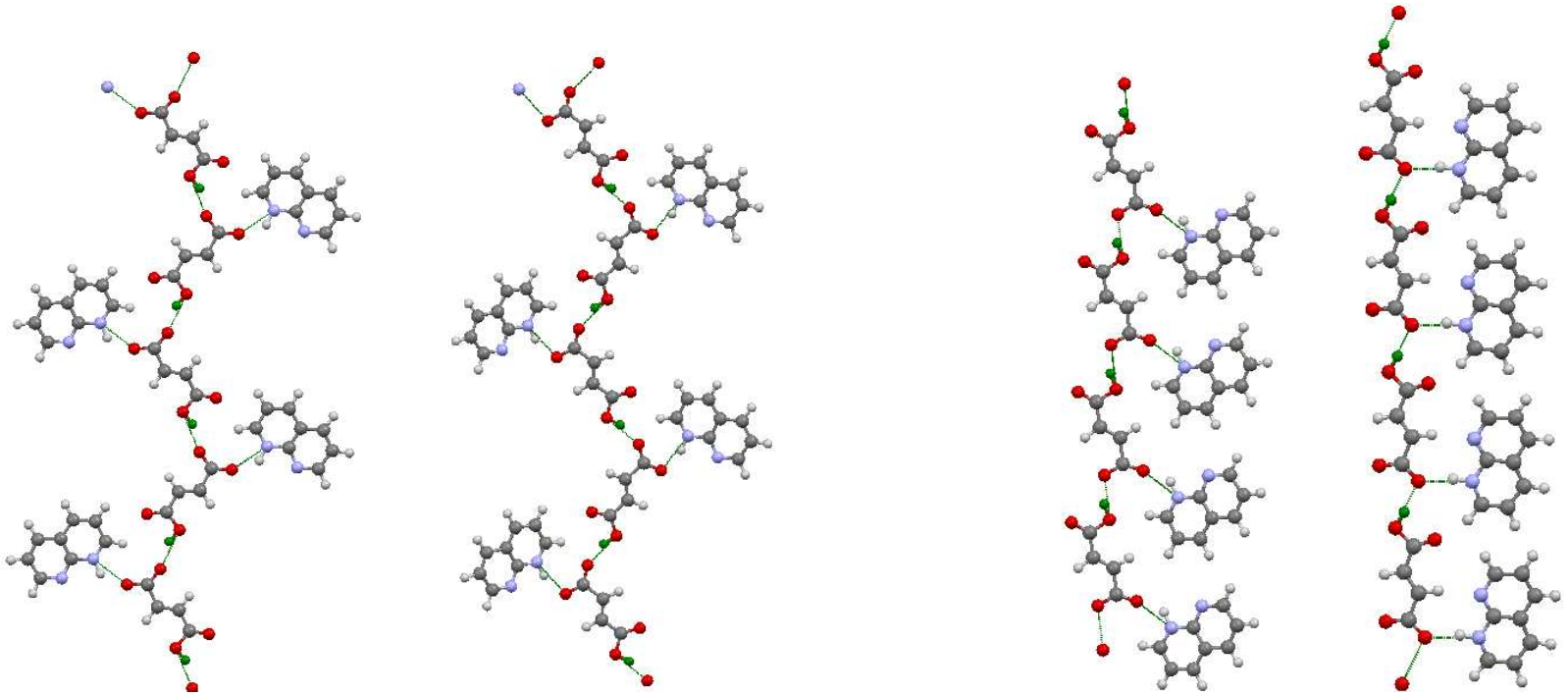
Figure 7.8: Rigid body crystal energy landscape of **XIX**. The predicted lattice energy minima are denoted according to the fumarate conformation used in the search and the graph set of the hydrogen bond motif (common packings illustrated in Scheme 7.1 and Scheme 7.2). The conformation of the fumarate ion is identified according to one of the following abbreviations: st1: s-*trans*(1), sc1: s-*cis*(1) and sc2: s-*cis*(2).

s-*trans*(2) fumarate conformation do not have lattice energies that are within 12 kJ mol^{-1} of the global minimum and as a consequence do not appear in Figure 7.8. From the figure, the predicted global minimum and the second ranked structures are based on the s-*cis(2)* conformation for the fumarate ion. This may be surprising given the high relative energy of the s-*cis*(2) conformation when compared to the other three fumarate conformations shown in Figure 7.7. However, the results in Figure 7.8 suggest that for these two structures, the favourable intermolecular interactions observed in the crystal are enough to compensate for the high intramolecular energy penalty of 13.08 kJ mol^{-1}. For the more stable s-*trans*(2) fumarate conformation with a relative intramolecular energy of 6.65 kJ mol^{-1}, the possible intermolecular interactions are such that there are no hypothetical structures with lattice energies that are within 12 kJ mol^{-1} of the global minimum. This balance between the inter- and intra-molecular energies may explain why the relatively high energy s-*cis*(2) fumarate conformation is observed in the CSD while the lower energy s-*trans*(2) conformation is not. One other general observation that can be made regarding the CSD structures with s-*cis*(2) fumarate conformation is

that the majority of these structures contain tertiary or quaternary ammonium counterions instead of aromatic cations such as the 1,8-naphthyridinium ion of **XIX**.

The difference in the lattice energies of the two most stable structures from the search, which consist of fumarate ions with s-*cis*(2) conformation, is 2.57 kJ mol^{-1}. If we look at the predicted structures with the s-*cis*(1) or s-*trans*(1) fumarate conformations, we find that they too have one energetically preferred structure in lattice energy. For structures with an s-*trans(1)* fumarate conformation, the most stable lattice energy minimum is the third ranked structure and this is 2.64 kJ mol^{-1} more stable than the nearest lattice energy minimum with the same fumarate conformation. For structures with the s-*cis(1)* fumarate conformation, there is a 3.64 kJ mol^{-1} energy difference between the two most stable structures. Thus, the crystal energy landscape shows one energetically preferred packing for each of these fumarate conformations.

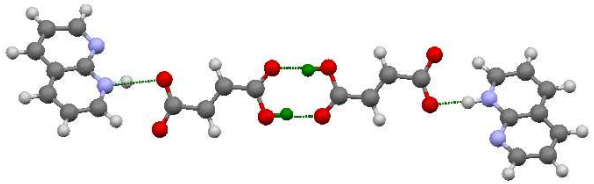
The crystal packings adopted by the predicted structures of **XIX** were compared and the four most common hydrogen bond motifs in the predicted crystal energy landscape are shown in Scheme 7.1 and Scheme 7.2. The most frequently occurring hydrogen bond motif in the rigid body crystal energy landscape (Figure 7.8) is ribbon motif 1 [graph set: $C_1^1(7)$-*a*]. This ribbon motif is observed in the hypothetical structures with s-*cis*(1) and s-*trans*(1) fumarate conformations and has an overall occurrence frequency of 41 % in the crystal energy landscape. The competing $R_2^2(8)$-*a* dimer motif (Scheme 7.2), which is also observed in structures with s-*cis*(1) and s-*trans*(1) fumarate conformations, has an overall occurrence frequency of 18 %. From Chapter 5, we saw that the $R_2^2(8)$ dimer motif was the most frequently occurring acid-acid interaction in the crystal energy landscape of the salt and cocrystal molecular structures of **IYUPEX**[186]. The $C_1^1(7)$-*b* chain motif illustrated in Scheme 7.1 has an overall occurrence frequency of 18 % and differs from the ribbon motif 1 network because the cations are stacked on the same side of the fumarate chain. The remaining $C_1^1(7)$-*c* ribbon motif 2 (Scheme 7.2) network has a similar one dimensional stacking of cations found in the $C_1^1(7)$-*b* chain motif, but the different s-*cis*(2) conformation of the fumarate ions leads to ribbons rather than chains. With the exception of the $R_2^2(8)$-*a* dimer motif, the remaining three types of packings shown in Scheme 7.1 and Scheme 7.2 are observed in the structures of fumarate salts retrieved from the CSD



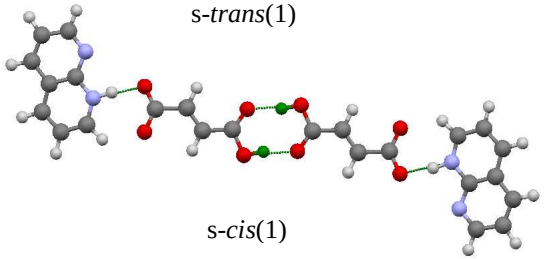
s-*trans*(1) s-*cis*(1) s-*trans*(1) s-*cis*(1)

	RIBBON MOTIF 1	**CHAIN MOTIF**
	$C_1^1(7)$-*a*	$C_1^1(7)$-*b*
% Occurrence of Motif*	41%	18%

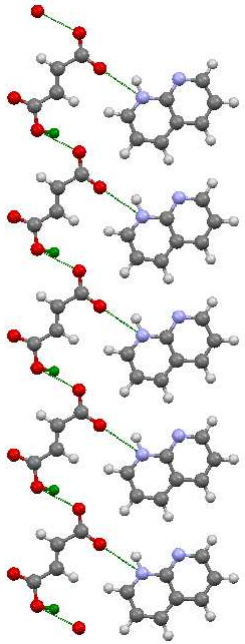
Scheme 7.1: Illustration of the ribbon motif 1 and chain motif hydrogen bonds found in structures with s-*cis*(1) or s-*trans*(1) fumarate conformation as observed in the predicted crystal energy landscape of **XIX**. Shown in green are the positions of the carboxylic acid proton on the monoionised fumarate ion. * For both motifs, the one dimensional stacking of the 1,8-naphthyridinium ions are sufficiently similar regardless of the fumarate conformation. Thus, the quoted % occurrence takes into account structures with the s-*cis*(1) and s-*trans*(1) fumarate conformations.



s-*trans*(1)



s-*cis*(1)



s-*cis*(2)

DIMER MOTIF

RIBBON MOTIF 2

$R_2^2(8)$-*a*

$C_1^1(7)$-*c*

% Occurrence of Motif:* 18%

18%

Scheme 7.2: Illustration of the dimer and ribbon motif 2 hydrogen bond networks observed in the predicted crystal structures of **XIX**. * The quoted % occurrence of the dimer motif is for structures with s-*cis*(1) or s-*trans*(1) fumarate conformation. The ribbon motif 2 network is only observed in structures with s-*cis*(2) fumarate conformation and the quoted % occurrence reflects this.

(appendix, Table 7.4). There are two structures that display an $R_2^2(8)$-*b* motif between the carboxylic acid and pyridine nitrogens on the 1,8-napthyridinium cation. One nitrogen atom acts as a hydrogen bond donor and the other a hydrogen bond acceptor. These structures are energetically unstable when compared to the other hypothetical structures.

7.3.4 Crystal energy landscape following refinement of flexible degrees of freedom

For each structure shown in Figure 7.8, CrystalOptimizer was used to re-minimise the lattice energy as a sum of an intermolecular and intramolecular components ($E_{latt}=U_{inter} + \Delta E_{intra}$). The resulting lattice energy landscape is shown in Figure 7.9 and unlike the rigid body crystal energy landscape (Figure 7.8) there doesn't appear to be a distinct energetically favoured structure. This is true if you focus on structures with a particular conformation of the fumarate ion or you look at the crystal energy landscape holistically. Thus small changes in molecular conformation can lead to improved hydrogen bond interactions between the ions and this has a significant effect in re-ranking the predicted structures in terms of lattice energy. Of the four lowest energy structures shown in Figure 7.9, the second and third ranked structures have the same crystal densities. They are however based on the s-*cis*(1) and s-*cis*(2) fumarate conformations with packings based on the chain or ribbon motif 2 networks respectively. The first and fourth ranked structures are based on the s-*cis*(1) and s-*trans*(1) fumarate conformations respectively. The majority of the low energy structures shown in Figure 7.9 have fumarate ions with s-*cis*(1) conformation. In fact, 13 of the 20 most stable structures have this fumarate conformation. By contrast, a survey of the CSD reported in Table 7.4 has shown that this conformation is found in roughly a third of all experimental fumarate salts surveyed. For each structure shown in Figure 7.9, an estimate of the Helmholtz Free energy (298 K) was made and the results are summarised in Table 7.3 of the appendix.

If we consider the crystal packings adopted by the lattice energy minima shown in Figure 7.9, we find that the most stable structures are based on the $C_1^1(7)$-*a* ribbon motif 1 packing (Scheme 7.1).

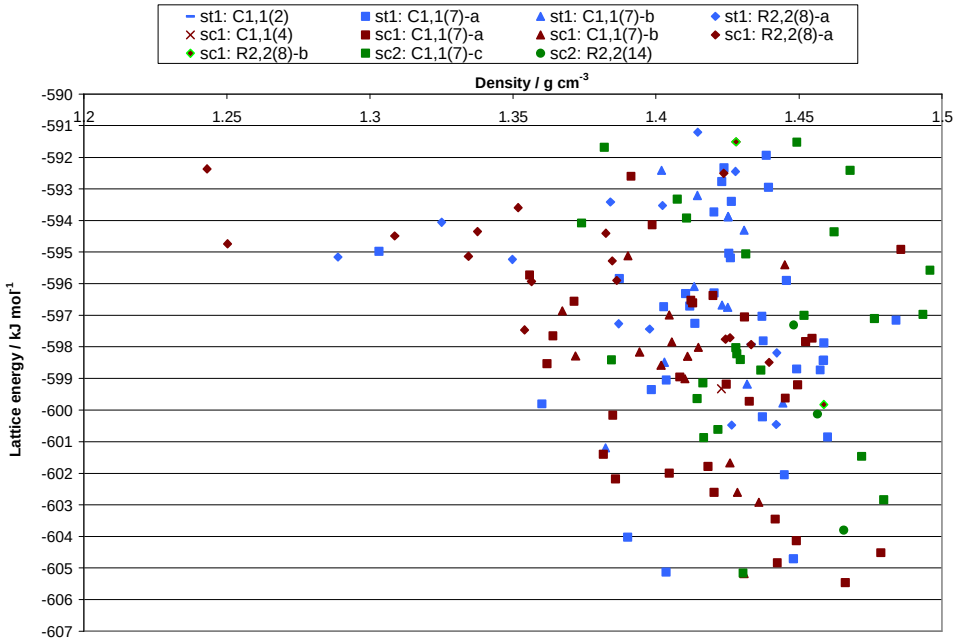
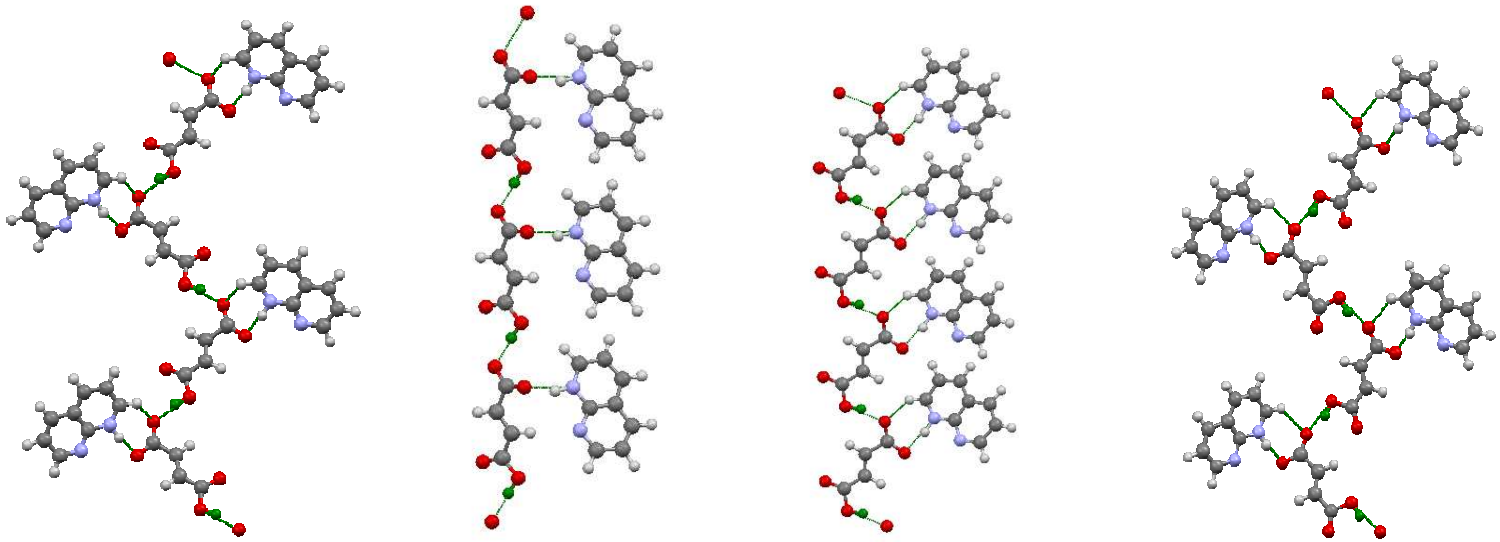


Figure 7.9: Crystal energy landscape after optimising the flexible torsional degrees of freedom on the fumarate ion. The different colours correspond to structures with different fumarate conformations and the shape of the symbol indicates the type of hydrogen bond motif observed in the crystal. The conformation of the fumarate ion is identified according to one of the following abbreviations: st1: s-*trans*(1), sc1: s-*cis*(1) and sc2: s-*cis*(2).

7.4 Choice of candidate structures

Under the rules of the CCDC, only three predicted structures may be submitted as 'guesses' for each experimental structure of a given blind test candidate. For small organic molecules, the results of previous blind tests[212,239,246,247] have shown that the choice of these structures is often complicated by the appearance of a crystal energy landscape with many energetically competitive packings. This thesis has shown that the rigid body crystal energy landscapes of salts can have one or two energetically favoured structures (Chapters 5 and 6). It is important to note that this is not a general observation for all salts, as the results of CSP and the details of the crystal energy landscape are system dependent. The extent, to which the low energy structures of a search correspond to the experimentally determined structure(s), will depend on the nucleation pathway taken by the growing crystallites. Thus the crystal structures of 4-dimethylaminopyridinium maleate (Chapter 5) and amantadine hydrochloride (Chapter 6) were predicted as the global minimum in lattice energy because they happened to crystallise as the only solid form or most stable solid form respectively. For **XIX**, the



	RANK 1	RANK 2	RANK 3	RANK 4
Structure ID	14_sc1	1_hs_sc1	635_sc2	7_st1
ΔE_{latt} / kJ mol^{-1}:	-605.46	-605.18	-605.16	-605.13
***A* (298 K) / kJ mol^{-1}:**	-617.90	-616.90	-616.00	-630.65
ρ / g cm^{-3}:	1.47	1.43	1.43	1.40

Scheme 7.3: Illustration of the hydrogen bond motifs and tabulation of the relative stabilities and crystal densities of the four most stable structures from Figure 7.9. In the 'Structure ID' row, the conformation of the fumarate ion is identified according to one of the following abbreviations: sc1: s-*cis*(1), sc2: s-*cis*(2) and st1: s-*trans*(1). For the rank 2 structure the extra "hs" term in the Structure ID shows that this structure was obtained from the CrystalPredictor search in high symmetry space groups. Where there is no "hs" in the Structure ID, this should be taken to mean that the structure was obtained from a CrystalPredictor search in the common space groups.

rigid body (Figure 7.8) crystal energy landscape does have one preferred structure in lattice energy. However, the flexible torsional degrees of freedom on the fumarate ion mean that we will be using the crystal energy landscape shown in Figure 7.9 when choosing the structures to be submitted to the CCDC.

We will consider the four lowest energy structures in Figure 7.9, since the majority of the low energy structures within 5 kJ mol^{-1} from the global minimum are based on different packings of the same hydrogen bond motifs found in these structures. The hydrogen bond motifs of the four lowest energy structures from Figure 7.9 are illustrated in Scheme 7.3. The second and third ranked structures in Figure 7.9 have the same cell densities but are based on fumarate ions with s-*cis*(1) and s-*cis*(2) conformations respectively. As a consequence the observed fumarate packings in the structures are different. All four structures have lattice energies that are within 0.40 kJ mol^{-1} and this makes the choice of the three submission structures difficult. Scheme 7.3 shows that the rank 2 structure is the only one that does not display an $R_2^2(7)$ dimer motif between the carboxylate and pyridinium ions and this is a consequence of the fumarate chain motif which prevents a head-on interaction between the oppositely charged ions. For all four structures, an estimate of the Helmholtz free energies (Chapter 2, section 2.5.4) at 298 K was made. The results, which are tabulated as part of Scheme 7.3 show that the rank 4 and rank 1 structures (using lattice energies) are the most stable and second most stable structures on the free energy surface. These structures have the s-*trans*(1) and s-*cis*(1) fumarate conformations respectively and are based on the $C_1^1(7)$-*a* ribbon motif 1 packing. These differences in conformation and the relative stability of the packings on the lattice energy and free energy surfaces make them good candidates for submission. These structures account for two of the three possible structures submitted to the CCDC. The final structure will either be the rank 2 or 3 structure. The estimates of the Helmholtz free energies (Scheme 7.3) has shown that the relative stabilities of the rank 2 and 3 structures widens to 0.90 kJ mol^{-1} and this is clearly an improvement over the 0.02 kJ mol^{-1} difference observed using the lattice energy. The more stable structure at 298 K is the rank 2 structure. However, if we were to neglect the rank 3 structure in favour of the rank 2 structure, we would be discarding the global minimum structure from the rigid body search (rank 3) and choosing a rank 2 structure with the same fumarate conformation as rank 1. The CSD survey reported in Table 7.4 shows that the s-*cis*(2) fumarate conformation is statistically as likely as the s-

cis(1) and s-*trans*(1) conformations given that each accounts for roughly a third of all fumarate ions surveyed. This provides a case for including the rank 3 structure among the three submitted to the CCDC. Thus, the three submitted structures will be those ranked 1, 3 and 4 on the basis of the lattice energy (Figure 7.9).

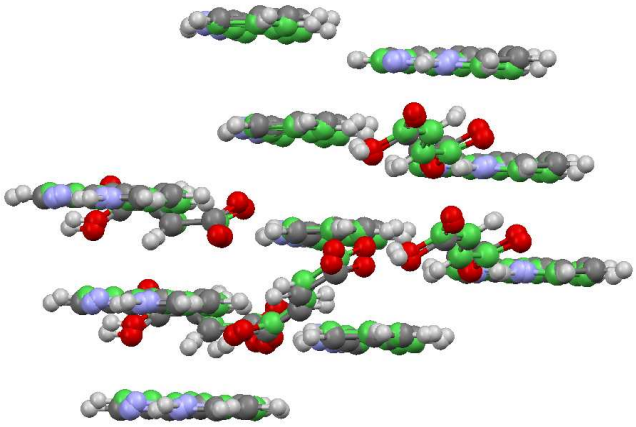
7.5 Discussion in light of the experimental structure

Table 7.2 gives the crystallographic parameters and lattice energies of the three submitted structures and for comparison, the same details for the experimental structure (Exp) and *ExpMinOpt* are given. If we compare the lattice parameters of the submitted global minimum structure (14_sc1) and those for the *ExpMinOpt* of **XIX**, we find similarities. However, the two structures are based on different orthorhombic space groups and the experimental structure crystallises with an s-*trans*(1) fumarate conformation whilst the predicted global minimum structure displays an s-*cis*(1) fumarate conformation. Despite these differences, both structures adopt the same $C_1^1(7)$-*a* fumarate ribbon (Scheme 7.1) motif and appear to have the same packing of the ions when the crystals are viewed in the *ac* plane (Figure 7.10). However, there are noticeable differences in the positions of both ions in 3-dimensions. The third submitted structure from the search, 7_st1, has the same fumarate conformation as the experimental structure but there is no local similarity in the $C_1^1(7)$-*a* fumarate ribbons or stacking of the cations. This difference in cation stacking is caused by the use of different carboxylate oxygens for the N^+-H$\cdots$O hydrogen bonds in the two structures.

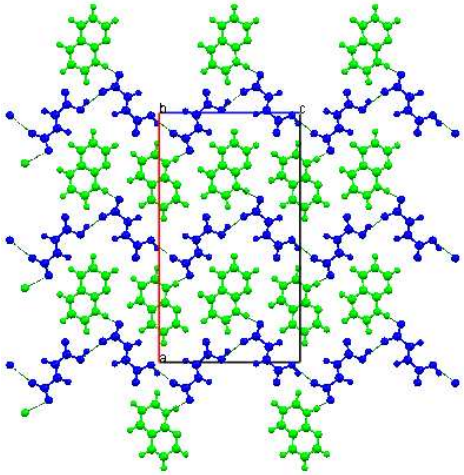
Comparison of the experimental structure with the list of extended structures (Table 7.3) showed that the experimental structure was not predicted in the search. This is despite the fact that 100,000 local minimisations were performed by CrystalPredictor in the statistically most common space groups (*cf.* section 7.2) - including the experimental $Pca2_1$ space group - for each conformation of the fumarate ion. The potential derived charge model used in the search ranked the experimental structure as 24 kJ mol^{-1} higher in energy than the most stable structure with the s-*trans*(1) fumarate conformation. Many of the more stable $Pca2_1$ structures were found at least 5 times and it seems that a more extensive search would have located the experimental structure. This was confirmed by repeating the search in just the $Pca2_1$ space group.

Structure ID	E_{latt} */ kJ mol^{-1}	ρ / g cm^{-3}	Space group	a / Å	b / Å	c / Å	α / °	β / °	γ / °
14_sc1	-605.461	1.466	$Pna2_1$	22.405	3.856	12.910	90	90	90
635_sc2	-605.162	1.430	P-1	8.980	6.438	12.243	87.642	56.874	101.343
7_st1	-605.127	1.404	Cc	10.577	9.585	12.260	90	69.626	90
Exp	-	1.481	$Pca2_1$	23.501	3.714	12.653	90	90	90
ExpMinOpt	-594.985	1.433	$Pca2_1$	22.976	3.787	13.120	90	90	90

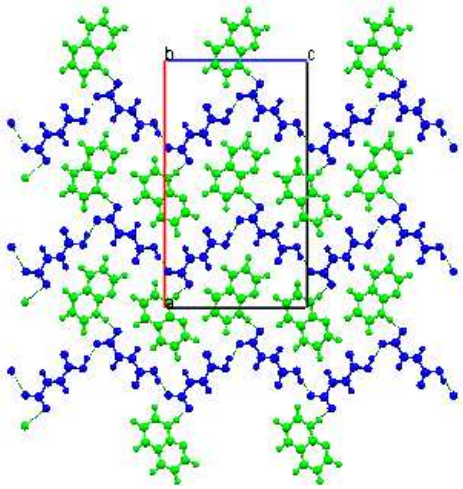
Table 7.2: Crystallographic lattice parameters and estimated lattice energies of the three submitted structures, the experimental structure and ExpMinOpt of **XIX**. * All quoted lattice energies are true after refining the flexible torsion angles of the fumarate ion with CrystalOptimizer.



(a)



(b)



(c)

Figure 7.10: Part (a) shows an overlay of the experimental (coloured green) and matching predicted structure (coloured by element) from the search. The $RMSD_{15}$ for the overlay is 0.265 Å. The matching structure was found retrospectively in a search performed after the announcement of the blind test results. Parts (b) and (c) show the crystal packing in the experimental and predicted global minimum structure from the search (Figure 7.9) respectively.

Despite the fact that the experimental structure was not predicted in the search, it is important to emphasise that the fumarate ribbons observed in the experimental structure are found in many of the predicted low energy structures such as the global minimum in lattice energy. Using a distributed multipole model for the lattice energy and refining the flexible torsional degrees of freedom of the experimental structure leads to a relative lattice energy estimate of 10.48 kJ mol^{-1} with respect to the predicted global minimum. Thus even if the experimental structure had been predicted in the search, it is too unstable to have been considered for submission on energetic grounds alone. The experimental structure of **XIX** is isostructural with the known quinolinium fumarate salt, **RABYID**. A search for hypothetical crystal structures of **RABYID**, performed prior to the work on **XIX** predicted the experimental structure to have a relative lattice energy of 13 kJ mol^{-1} with respect to the global minimum. Thus, the parallel between the modelling results for both salts and the fact that they are isostructural suggests that a model which reproduces the observed structure of **RABYID** at or close to the global minimum in lattice energy may also prove successful for **XIX** and vice versa.

When the molecular structures of the ions in **XIX** are compared with those of the previous salts modelled in this thesis, one obvious difference is the number of possible hydrogen bond interaction sites as we find that both ions have one hydrogen bond donor and acceptor functional groups. The salt minor product of **IYUPEX** (Chapter 5) is an example of another system with more than one functional group on the anion that is capable of intermolecular hydrogen bonding. What makes **IYUPEX** less complicated than **XIX** is that the pyridinium ion only has an N^{+}-H functional group for hydrogen bonding in contrast to the 1,8-naphthyridinium cation of **XIX** which also has a pyridine N acceptor. This creates a greater number of plausible hydrogen bond interactions in **XIX**. As a consequence, a correct ranking of the relative energies of the predicted structures is more demanding of a force field that can capture the effects of many competing intermolecular interactions. Whilst the FIT(H_O,H_N) force field has proven adequate for salt crystal structures containing one dominant hydrogen bond interaction between the cation and anion (Chapters 5 and 6), the results in this chapter show that it is not suitable for modelling systems with more than one hydrogen bond functional group on one (**RABYID**) or both (**XIX**) of the ions. Thus future work on salts with mono-ionised dicarboxylic acids should move towards an *ab initio* based model for the lattice energy that treats the induction contribution explicitly.

An alternative explanation for the high relative energy of **XIX** is that the crystal structure determined is a kinetic product of crystallisation rather than the thermodynamic polymorph assumed in the modelling. Whilst this cannot be ruled out without a thorough screen, the inadequacy of the intermolecular force field should be addressed prior to initiating an extensive polymorph screen. If **XIX** is indeed polymorphic, the results from the crystal energy landscape suggest that the other polymorphs are unlikely to have crystal packings that deviate from a fumarate ribbon motif, since this is a frequently observed feature among many of the predicted low energy structures. However, the stacking of the cations may differ as the N^+-H donor on the cation may form hydrogen bonds with either of the two carboxylate oxygens. The fact that the free base, 1,8-naphthyridine, is not commercially available suggests that a synthetic procedure may have been used to modify a closely related analog in order to obtain sufficient quantities of the base for crystallisation with fumaric acid. If this was the case, the effects of impurities should be investigated as they are known to affect the crystallisation pathway and hence the observed crystal form.

Of the fifteen groups participating in the blind test, only two groups included the experimental structure of **XIX** among the three submitted structures. Van Eijck used the UPACK[200] package and an empirical Buckingham potential in conjunction with a distributed multipole model derived from the dimer. This differs from the approach used in this thesis which used the isolated gas phase conformational minima of the ions to derive the intramolecular energies and multipole models. Van Eijck predicted the experimental structure of **XIX** as the second ranked structure from his search. Kendrick predicted the experimental structure of **XIX** as rank 20 using the GRACE[140] package which implements a dispersion-corrected density functional theory method[139] to perform the final stability ranking of structures. The decision to include the rank 20 structure among the three submitted structures was made on the basis of its isostructurality to **RABYID**, which had also proven to be energetically unfavourable in a search performed prior to the modelling of **XIX**. Only two participants (Day and Boerrigter) had the experimental structure of **XIX** in the extended list of plausible structures submitted with their chosen three. The 1,8-naphthyridinium fumarate salt, **XIX**, was the second least attempted system by participants of the blind test. This indicates the level of confidence some participants felt in adapting their method of crystal structure prediction to the modelling of organic salt crystal structures.

7.6 Conclusions

Reported in this chapter are the results of a blind crystal structure prediction for the 1,8-naphthyridinium fumarate salt, **XIX**. A survey of the Cambridge Structural Database (CSD) revealed 3 distinct conformations [s-*cis*(1), s-*cis*(2) and s-*trans*(1)] of the fumarate ion differing in the torsions defining rotation about the OH and COOH groups. A fourth conformation [s-*trans*(2)] absent from the fumarate salts surveyed in the CSD was found to be a conformational minimum of the fumarate ion. All four conformations of the fumarate ion were used in separate rigid body searches for hypothetical crystal structures of **XIX** and in all cases the same gas phase optimised conformation of the 1,8-naphthyridinium ion was used. The results of the rigid body search for hypothetical crystal structures of **XIX** revealed a global minimum structure based on the highest energy s-*cis*(2) fumarate conformation. All rigid body structures within 12 kJ mol^{-1} of the global minimum were refined by optimising the flexible torsional degrees of freedom on the fumarate ion and re-minimising the lattice energy as a sum of an intermolecular and intramolecular component. This led to a crystal energy landscape with many energetically competitive structures but the majority of the low energy structures were based on the $C_1^1(7)$-*a* ribbon motif 1 packing. The three lowest energy structures based on the three different fumarate conformations observed in the CSD were submitted as plausible candidates for the blind test challenge. The experimental structure was later found to have the same fumarate ribbon and cation stacking observed in the predicted global minimum but different fumarate conformation and extended packing of the ions.

An exact match for the experimental structure of **XIX** was not found among the extended list of structures output by CrystalPredictor and this was rationalised on the basis of an insufficient search for hypothetical crystal structures of the salt. In CrystalPredictor, the extent to which a search is complete is related to the number of times the same structure is found by the minimiser. If the same structure is found many times, this is an indication that the search is either complete or nearing completion. A preliminary search for hypothetical crystal structures of quinolinium fumarate, **RABYID**, was successful in finding the experimental structure with 100,000 local minimisations in the statistically most common space groups. For **XIX**, the more stable structures in the experimental $Pca2_1$ space group were found many times, but the high

relative energy of the experimental structure precluded it from being found with just 100,000 local minimisations.

Overall, the first salt system posed as a blind test candidate has proven to be a challenge. The fact that the majority of methods used by participants predicted the experimental structure of **XIX** too high in relative energy, suggests that better models for the relative energies of predicted structures are needed for salts with many competing intermolecular interactions. The success of van Eijck in predicting the experimental structure as rank 2 suggests that calculating the atomic multipole moments from the gas phase dimer may lead to more realistic relative energies for salts. This is because the calculated multipole moments have absorbed the effects of polarisation between the hydrogen bonded ions. At present, DMACRYS uses the atomic multipole moments derived from the isolated wavefunctions of the individual ions. Thus future work on salts should consider the use of the dimer wavefunction, especially when there are many hydrogen bond functional groups on the molecular ions.

7.7 Appendix

7.7.1 Predicted most stable lattice energy minima of 1,8-naphthyridinium fumarate XIX

Structure ID	E_{latt}/ kJ mol^{-1}	A / kJ mol^{-1}	ρ / g cm^{-3}	Space Group	a / Å	b / Å	c / Å	α / °	β / °	γ / °	Graph Set
14_sc1	-605.461	-617.896	1.466	Pna21	22.405	3.856	12.910	90	90	90	$C_1^1(7)$ -a
1_hs_sc1	-605.176	-616.901	1.431	Fdd2	28.096	20.238	8.041	90	90	90	$C_1^1(7)$ -b
635_sc2	-605.162	-616.004	1.430	P-1	8.980	6.438	12.243	87.642	56.874	101.342	$C_1^1(7)$ -c
7_st1	-605.127	-630.646	1.404	Cc	10.577	9.585	12.260	90	69.626	90	$C_1^1(7)$ -a
3_sc1	-604.831	-617.280	1.442	Pbca	15.527	12.900	11.321	90	90	90	$C_1^1(7)$ -a
2_hs_st1	-604.711	-617.649	1.448	Pccn	11.994	13.589	13.859	90	90	90	$C_1^1(7)$ -a
3_hs_st1	-604.710	-617.645	1.448	Pccn	11.994	13.590	13.858	90	90	90	$C_1^1(7)$ -a
2_hs_sc1	-604.517	-615.998	1.479	Pc	3.815	11.481	12.625	90	90.135	90	$C_1^1(7)$ -a
69_sc1	-604.132	-615.639	1.449	Pbca	20.946	7.713	13.971	90	90	90	$C_1^1(7)$ -a
116_st1	-604.021	-616.305	1.390	P21/c	10.649	8.363	13.211	90	89.609	90	$C_1^1(7)$ -a
2_sc2	-603.800	-615.963	1.466	P21/a	14.850	6.754	11.150	90	86.219	90	$R_2^2(14)$
15_sc1	-603.452	-616.393	1.442	I2/c	14.828	11.242	13.718	90	82.849	90	$C_1^1(7)$ -a
30_sc1	-602.917	-615.004	1.436	P21/c	8.068	10.096	14.758	90	71.340	90	$C_1^1(7)$ -b
1_sc2	-602.844	-614.306	1.480	P-1	9.236	10.358	6.355	66.905	98.392	95.807	$C_1^1(7)$ -c
5_sc1	-602.607	-616.029	1.420	P21/c	7.725	11.548	14.230	90	65.098	90	$C_1^1(7)$ -a
25_sc1	-602.601	-614.284	1.428	P-1	10.113	7.699	8.110	108.974	90.216	74.318	$C_1^1(7)$ -b
10_sc1	-602.177	-614.782	1.386	Cc	10.237	10.138	12.083	90	70.228	90	$C_1^1(7)$ -a

39_st1	-602.048	-613.846	1.445	P212121	7.445	10.972	13.855	90	90	90	$C_1^1(7)$ -a
9_sc1	-602.000	-615.154	1.405	Pbca	11.453	15.687	12.960	90	90	90	$C_1^1(7)$ -a
65_sc1	-601.787	-613.515	1.418	P212121	8.047	10.921	13.122	90	90	90	$C_1^1(7)$ -a
23_sc1	-601.674	-613.415	1.426	P-1	7.923	10.523	8.964	64.739	121.948	102.803	$C_1^1(7)$ -b
8_sc2	-601.467	-612.764	1.472	P21/c	10.021	9.811	12.122	90	68.793	90	$C_1^1(7)$ -c
109_sc1	-601.398	-613.667	1.382	Pna21	13.358	8.625	10.274	90	90	90	$C_1^1(7)$ -a
62_st1	-601.200	-613.301	1.382	P21/c	7.493	8.030	20.466	90	106.098	90	$C_1^1(7)$ -b
11_sc2	-600.875	-612.937	1.417	Pna21	12.352	11.201	8.344	90	90	90	$C_1^1(7)$ -c
1_st1	-600.856	-613.741	1.460	Pbca	11.951	14.336	13.077	90	90	90	$C_1^1(7)$ -a
9_sc2	-600.619	-611.722	1.422	P21/c	6.493	8.242	22.317	90	105.605	90	$C_1^1(7)$ -c
259_st1	-600.479	-612.385	1.426	P21/n	10.533	10.052	10.899	90	83.451	90	$R_2^2(8)$ -a
53_st1	-600.456	-612.359	1.442	P-1	7.163	8.236	10.556	109.694	97.888	98.932	$R_2^2(8)$ -a
5_st1	-600.216	-613.365	1.437	P21/c	7.367	13.026	13.766	90	59.471	90	$C_1^1(7)$ -a
146_sc1	-600.163	-612.272	1.385	P21/c	11.282	11.848	9.563	90	112.518	90	$C_1^1(7)$ -a
32_sc2	-600.131	-612.364	1.456	C2/c	26.101	6.743	14.876	90	59.069	90	$R_2^2(14)$
69_hs_sc1	-599.821	-610.507	1.459	Pn	11.343	8.485	5.867	90	83.071	90	$R_2^2(8)$ -b
38_st1	-599.809	-612.641	1.360	P21/c	8.790	11.055	12.580	90	100.403	90	$C_1^1(7)$ -a
118_st1	-599.777	-611.618	1.444	P-1	8.932	10.708	7.593	116.172	98.573	60.669	$C_1^1(7)$ -b
5_sc2	-599.744	-623.502	1.500	P1	7.132	6.598	7.653	88.675	119.520	65.303	$C_1^1(7)$ -c
28_sc1	-599.722	-612.074	1.433	C2/c	7.733	13.433	21.997	90	87.611	90	$C_1^1(7)$ -a
14_sc2	-599.635	-612.158	1.414	P212121	8.056	25.617	5.603	90	90	90	$C_1^1(7)$ -c
8_sc1	-599.622	-	1.445	P21/c	7.596	21.387	7.071	90	80.088	90	$C_1^1(7)$ -a
16_st1	-599.355	-611.971	1.398	Pna21	13.620	8.590	9.996	90	90	90	$C_1^1(7)$ -a

16_sc1	-599.324	-611.660	1.423	P212121	5.076	15.957	14.191	90	90	90	$C_1^1(4)$
38_sc1	-599.200	-612.898	1.450	P21	3.805	14.810	10.250	90	77.566	90	$C_1^1(7)$ -a
21_sc1	-599.181	-611.846	1.425	P21/n	7.611	12.471	12.096	90	89.255	90	$C_1^1(7)$ -a
37_st1	-599.178	-611.466	1.432	P21/c	7.430	8.201	18.747	90	89.405	90	$C_1^1(7)$ -b
118_sc2	-599.144	-610.743	1.416	P21/c	6.415	18.597	10.425	90	111.804	90	$C_1^1(7)$ -c
6_hs_st1	-599.047	-612.171	1.404	Pccn	14.057	11.731	14.132	90	90	90	$C_1^1(7)$ -a
236_sc1	-599.009	-610.946	1.410	C2/c	20.967	8.176	15.006	90	115.609	90	$C_1^1(7)$ -b
1_sc1	-598.955	-612.042	1.408	P212121	14.380	8.577	9.414	90	90	90	$C_1^1(7)$ -a
4_sc1	-598.927	-611.657	1.414	P21/n	10.641	10.726	10.136	90	90.210	90	$R_2^2(8)$ -a
82_sc2	-598.739	-610.734	1.437	P21/c	6.298	14.205	12.933	90	79.709	90	$C_1^1(7)$ -c
32_st1	-598.733	-610.973	1.457	Ia	6.733	23.267	7.383	90	75.961	90	$C_1^1(7)$ -a
2_st1	-598.702	-611.729	1.449	I2/c	14.360	11.703	13.439	90	87.839	90	$C_1^1(7)$ -a
57_sc1	-598.578	-610.611	1.402	P21/c	8.504	9.799	16.573	90	122.353	90	$C_1^1(7)$ -b
12_sc1	-598.529	-610.677	1.362	Cc	11.614	8.697	12.711	90	69.280	90	$C_1^1(7)$ -a
19_sc1	-598.490	-611.076	1.440	P-1	8.071	7.662	9.953	103.758	73.837	85.994	$R_2^2(8)$ -a
200_st1	-598.484	-610.726	1.403	P21/c	7.493	8.317	19.588	90	72.750	90	$C_1^1(7)$ -b
9_st1	-598.434	-611.546	1.458	P21/c	6.578	13.109	13.657	90	72.211	90	$C_1^1(7)$ -a
4_st1	-598.429	-611.540	1.459	P21/c	6.577	13.108	13.658	90	107.785	90	$C_1^1(7)$ -a
56_sc2	-598.415	-610.418	1.384	P21	10.438	7.332	7.976	90	75.371	90	$C_1^1(7)$ -c
2_hs_sc2	-598.402	-617.510	1.429	P2/n	12.336	6.430	14.461	90	85.907	90	$C_1^1(7)$ -c
933_sc1	-598.297	-610.322	1.411	P-1	11.211	7.656	9.089	121.409	118.231	82.509	$C_1^1(7)$ -b
208_sc1	-598.290	-610.335	1.372	P21/c	10.334	8.113	15.103	90	109.700	90	$C_1^1(7)$ -b
6_sc2	-598.222	-610.830	1.428	I2/a	18.298	6.442	19.468	90	86.259	90	$C_1^1(7)$ -c

34_st1	-598.194	-610.409	1.442	P-1	9.747	7.692	8.318	97.054	74.313	74.907	$R_2^2(8)$ -a
44_sc1	-598.160	-610.446	1.394	P21/c	8.157	10.061	15.444	90	112.270	90	$C_1^1(7)$ -b
76_sc2	-598.030	-609.443	1.428	P21/a	14.733	12.331	6.451	90	77.754	90	$C_1^1(7)$ -c
393_sc1	-598.015	-610.259	1.415	A2/n	19.627	8.222	15.022	90	72.482	90	$C_1^1(7)$ -b
63_sc1	-597.926	-610.247	1.433	P-1	10.452	7.218	8.278	98.571	71.103	81.161	$R_2^2(8)$ -a
36_st1	-597.919	-611.607	1.433	P212121	15.646	4.115	17.722	90	90	90	C1,1(2)
31_st1	-597.877	-610.724	1.459	Pbcn	13.117	13.524	12.640	90	90	90	$C_1^1(7)$ -a
40_sc1	-597.843	-609.907	1.406	P21	7.688	9.981	8.158	90	68.338	90	$C_1^1(7)$ -b
35_sc1	-597.831	-609.783	1.452	Cc	3.829	21.068	15.340	90	65.509	90	$C_1^1(7)$ -a
18_st1	-597.811	-610.822	1.438	Pbcn	13.721	11.694	14.181	90	90	90	$C_1^1(7)$ -a
160_sc1	-597.758	-610.274	1.424	P21/n	11.056	9.970	10.547	90	81.012	90	$R_2^2(8)$ -a
31_sc1	-597.729	-609.836	1.455	Cc	3.822	21.086	14.186	90	79.578	90	$C_1^1(7)$ -a
11_sc1	-597.711	-609.910	1.426	P-1	8.015	11.840	7.662	116.214	92.920	113.745	$R_2^2(8)$ -a
2_sc1	-597.649	-616.418	1.364	Pna21	11.625	6.363	16.210	90	90	90	$C_1^1(7)$ -a
13_sc2	-597.524	-608.758	1.503	P-1	13.134	7.590	6.587	96.132	103.595	118.166	$C_1^1(7)$ -c
243_sc1	-597.463	-610.814	1.354	P21/c	6.799	11.188	16.888	90	70.075	90	$R_2^2(8)$ -a
292_st1	-597.432	-610.235	1.398	P21/c	7.911	7.082	22.157	90	109.512	90	$R_2^2(8)$ -a
7_sc2	-597.310	-609.490	1.448	P21/c	12.696	6.804	15.043	90	119.644	90	$R_2^2(14)$
78_st1	-597.273	-609.468	1.387	P21/n	11.211	10.544	9.977	90	90.975	90	$R_2^2(8)$ -a
40_st1	-597.261	-609.042	1.414	P21/n	10.927	10.317	10.288	90	85.978	90	$C_1^1(7)$ -a
29_st1	-597.158	-610.517	1.484	Pna21	20.611	3.749	14.262	90	90	90	$C_1^1(7)$ -a
3_sc2	-597.104	-	1.476	P-1	11.396	8.075	6.423	97.530	105.748	98.310	$C_1^1(7)$ -c
60_sc1	-597.060	-609.501	1.431	P21/c	7.331	10.191	16.539	90	67.648	90	$C_1^1(7)$ -a
13_st1	-597.032	-610.461	1.437	Pca21	11.950	6.614	14.399	90	90	90	$C_1^1(7)$ -a

4_sc2	-597.008	-608.919	1.452	Cc	10.556	10.316	10.655	90	76.132	90	$C_1^1(7)$ -c
103_sc1	-596.996	-609.063	1.405	P-1	7.994	8.200	10.084	91.189	106.877	111.574	$C_1^1(7)$ -b
34_sc2	-596.972	-608.797	1.493	P-1	13.404	7.060	6.524	67.806	75.649	92.313	$C_1^1(7)$ -c
135_sc1	-596.858	-609.115	1.367	P21/c	8.826	19.206	8.587	90	55.257	90	$C_1^1(7)$ -b
494_st1	-596.747	-611.073	1.425	C2/c	26.307	7.541	12.641	90	66.236	90	$C_1^1(7)$ -b
27_st1	-596.731	-608.796	1.403	P21/c	7.712	10.360	14.593	90	89.851	90	$C_1^1(7)$ -a
8_st1	-596.713	-609.236	1.412	P212121	8.653	13.008	10.291	90	90	90	$C_1^1(7)$ -a
28_st1	-596.678	-610.246	1.423	C2/c	26.499	7.537	12.638	90	65.574	90	$C_1^1(7)$ -b
161_sc1	-596.613	-608.567	1.413	P21/a	7.329	12.928	12.562	90	76.548	90	$C_1^1(7)$ -a
36_sc1	-596.559	-	1.371	P212121	10.358	8.677	13.269	90	90	90	$C_1^1(7)$ -a
7_sc1	-596.522	-609.223	1.412	Pna21	14.700	8.291	9.502	90	90	90	$C_1^1(7)$ -a
80_sc1	-596.378	-607.931	1.420	Pbca	7.356	24.340	12.866	90	90	90	$C_1^1(7)$ -a
4_hs_st1	-596.311	-	1.410	Iba2	13.920	11.764	14.162	90	90	90	$C_1^1(7)$ -a
49_st1	-596.297	-	1.420	C2/c	7.507	13.058	26.218	90	63.635	90	$C_1^1(7)$ -a
35_st1	-596.088	-	1.413	C2/c	12.756	7.569	24.015	90	93.460	90	$C_1^1(7)$ -b
423_sc1	-595.933	-615.905	1.356	P21/c	6.808	15.523	12.914	90	117.954	90	$R_2^2(8)$ -a
33_st1	-595.901	-607.742	1.446	P21/c	7.090	11.541	16.131	90	121.012	90	$C_1^1(7)$ -a
397_sc1	-595.898	-	1.386	P21/c	4.401	15.882	17.619	90	73.341	90	$R_2^2(8)$ -a
41_st1	-595.843	-608.559	1.387	P21/c	7.799	15.360	10.471	90	109.980	90	$C_1^1(7)$ -a
17_sc1	-595.726	-609.699	1.356	Pbca	16.259	12.983	11.428	90	90	90	$C_1^1(7)$ -a
36_sc2	-595.571	-608.050	1.496	P21/c	6.492	25.743	7.162	90	65.987	90	$C_1^1(7)$ -c
484_sc1	-595.405	-606.609	1.445	P-1	8.086	8.923	9.099	63.707	99.625	105.658	$C_1^1(7)$ -b
235_sc1	-595.280	-608.713	1.385	P21/c	4.411	15.853	17.577	90	73.915	90	$R_2^2(8)$ -a

811_st1	-595.230	-	1.350	P21/c	6.803	11.326	16.769	90	69.672	90	$R_2^2(8)$ -a
23_st1	-595.183	-	1.426	Pbca	14.711	13.177	11.832	90	90	90	$C_1^1(7)$ -a
544_st1	-595.158	-608.117	1.289	C2/c	28.050	7.344	12.966	90	108.188	90	$R_2^2(8)$ -a
81_hs_sc1	-595.133	-608.252	1.334	Fdd2	31.572	22.266	6.974	90	90	90	$R_2^2(8)$ -a
142_sc1	-595.115	-607.029	1.390	P-1	8.267	10.019	8.375	66.001	111.511	101.938	$C_1^1(7)$ -b
72_sc2	-595.058	-	1.431	C2/c	18.649	6.338	19.628	90	99.935	90	$C_1^1(7)$ -c
14_st1	-595.038	-608.461	1.425	Pbca	14.256	13.466	11.953	90	90	90	$C_1^1(7)$ -a
40_hs_st1	-594.981	-607.961	1.303	I41/a	13.684	13.684	26.806	90	90	90	$C_1^1(7)$ -a
39_sc1	-594.917	-607.851	1.486	Pna21	22.668	3.710	13.089	90	90	90	$C_1^1(7)$ -a
758_sc1	-594.742	-607.726	1.250	C2/c	28.247	7.083	13.118	90	94.566	90	$R_2^2(8)$ -a
440_sc1	-594.489	-608.045	1.309	I2/c	26.472	7.337	12.967	90	82.937	90	$R_2^2(8)$ -a
56_sc1	-594.404	-607.303	1.382	P21/a	14.987	11.324	6.986	90	86.190	90	$R_2^2(8)$ -a
18_sc2	-594.362	-606.535	1.462	P21/a	12.702	13.828	6.517	90	77.715	90	$C_1^1(7)$ -c
461_sc1	-594.348	-607.122	1.338	Pbca	19.811	9.824	12.564	90	90	90	$R_2^2(8)$ -a
51_st1	-594.307	-608.376	1.431	Pbca	15.053	24.074	6.308	90	90	90	$C_1^1(7)$ -b
51_sc1	-594.131	-607.592	1.399	P212121	14.940	7.004	11.173	90	90	90	$C_1^1(7)$ -a
378_sc2	-594.083	-605.628	1.374	P-1	11.142	10.761	6.378	81.488	68.513	56.948	$C_1^1(7)$ -c
40_sc2	-594.078	-	1.501	P-1	9.058	9.650	6.483	77.484	80.233	90.099	$C_1^1(7)$ -c
77_hs_st1	-594.057	-607.280	1.325	Fdd2	31.165	22.792	6.950	90	90	90	$R_2^2(8)$ -a
48_sc2	-593.920	-604.956	1.411	P-1	13.566	10.486	6.352	59.092	55.743	53.738	$C_1^1(7)$ -c
5_hs_st1	-593.875	-609.011	1.425	P2/c	6.258	7.537	24.329	90	89.997	90	$C_1^1(7)$ -b
54_st1	-593.874	-609.010	1.425	P21212	6.258	24.329	7.537	90	90	90	$C_1^1(7)$ -b
17_st1	-593.737	-606.897	1.420	P21/c	7.571	11.983	14.734	90	120.518	90	$C_1^1(7)$ -a

392_sc1	-593.592	-606.702	1.352	C2/c	25.704	7.378	13.519	90	109.301	90	$R_2^2(8)$ -a
280_st1	-593.529	-605.259	1.402	P21/n	11.262	9.530	11.395	90	72.478	90	$R_2^2(8)$ -a
360_st1	-593.418	-606.723	1.384	P21/c	4.381	15.806	17.614	90	104.366	90	$R_2^2(8)$ -a
56_st1	-593.396	-606.259	1.426	P21/n	7.325	12.262	13.032	90	78.396	90	$C_1^1(7)$ -a
59_sc2	-593.321	-604.798	1.407	P-1	6.503	10.224	9.162	76.905	99.918	86.317	$C_1^1(7)$ -c
42_st1	-593.204	-605.767	1.415	P21/c	7.494	9.371	16.525	90	94.929	90	$C_1^1(7)$ -b
11_st1	-592.956	-611.435	1.439	P212121	6.253	12.225	14.864	90	90	90	$C_1^1(7)$ -a
44_st1	-592.773	-	1.423	Aa	13.933	13.078	6.307	90	89.983	90	$C_1^1(7)$ -a
119_sc1	-592.601	-605.173	1.391	P21/c	8.022	11.211	15.091	90	119.993	90	$C_1^1(7)$ -a
744_sc1	-592.507	-605.074	1.424	P-1	7.153	11.818	8.944	69.013	81.693	54.722	$R_2^2(8)$ -a
671_st1	-592.448	-604.864	1.428	C2/c	20.690	7.175	16.705	90	112.527	90	$R_2^2(8)$ -a
43_st1	-592.414	-607.254	1.402	P21/c	7.564	23.475	7.604	90	59.771	90	$C_1^1(7)$ -b
27_sc2	-592.411	-604.964	1.468	P21/c	6.555	7.065	25.060	90	73.737	90	$C_1^1(7)$ -c
632_sc1	-592.368	-606.581	1.243	P21/a	13.181	6.967	14.329	90	88.98412	90	$R_2^2(8)$ -a
26_st1	-592.328	-605.669	1.424	P21	6.495	13.889	6.583	90	104.729	90	$C_1^1(7)$ -a
19_st1	-591.939	-610.432	1.439	P212121	13.118	6.245	13.877	90	90	90	$C_1^1(7)$ -a
224_sc2	-591.685	-604.479	1.382	P21/n	10.288	18.292	6.481	90	76.013	90	$C_1^1(7)$ -c
42_sc2	-591.528	-603.257	1.449	P-1	9.534	6.527	9.375	81.543	93.041	78.882	$C_1^1(7)$ -c
496_sc1	-591.507	-603.270	1.428	P-1	6.966	9.371	9.304	89.589	71.578	83.835	$R_2^2(8)$ -b
401_st1	-591.202	-604.106	1.415	P-1	7.131	9.642	8.923	72.678	81.361	90.681	$R_2^2(8)$ -a

Table 7.3: Tabulated predicted lattice energy minima of the most stable structures following refinement of the rigid body structures with respect to the flexible torsional degrees of freedom on the fumarate ion. The refinement of the rigid body structures was performed with CrystalOptimizer and the predicted structures are ranked according to the lattice energy (E_{latt}). For comparison, the estimated Helmholtz Free Energies at 298 K (A) are also given in the table. The structure

identification contains implicit information about the fumarate conformation: sc1=s-*cis*(1), sc2=s-*cis*(2) and st1=s-*trans*(1). Where the structure identification also contains "hs", this should be taken to mean that the structure was found following a search in high symmetry (hs) space groups. Otherwise the structure was found in a standard CrystalPredictor search involving the most common space groups for organic molecules.

7.7.2 Fumarate salts retrieved from the Cambridge Structural Database

	FUMARATE CONFORMATION		GRAPH SET OF FUMARATE MOTIF		TORSION ANGLE / °			
REFCODE	**FUMARATE 1**	**FUMARATE 2**	**FUMARATE 1**	**FUMARATE 2**	**FUMARATE 1 [θ_1 - OH]**	**FUMARATE 1 [θ_2 - COOH]**	**FUMARATE 2 [θ_1 - OH]**	**FUMARATE 2 [θ_2 - COOH]**
ABEJUN	s-*trans*(1)	-	$C_1^1(7)$ - RIBBON 1	-	-174.506	-178.863	-	-
AKETOZ	s-*cis*(2)	s-*cis*(2)	$C_1^1(7)$ - RIBBON 2	$C_1^1(7)$ - RIBBON 2	-0.114	-11.241	-0.074	-9.868
AMAVEP	s-*trans*(1)	s-*cis*(2)	$C_2^2(14)$	$C_2^2(14)$	166.701	165.175	1.524	0.726
BAHLEC	s-*cis*(2)	-	$C_1^1(7)$ - RIBBON 2	-	0	0	-	-
CLEMAS	s-*cis*(1)	-	$C_1^1(7)$ - RIBBON 1	-	169.963	-0.189	-	-
EKIKIS	s-*trans*(1)	s-*cis*(1)	$D_1^1(2)$	$D_1^1(2)$	-175.93	179.584	-164.03	0.21
EWELUN	s-*cis*(1)	-	$C_1^1(7)$ - RIBBON 1	-	164.818	-0.986	-	-
HUSSUJ	s-*cis*(1)	-	$C_1^1(7)$ - CHAIN	-	179.867	-5.52	-	-
HUSTAQ01	s-*cis*(2)	-	$C_1^1(7)$ - RIBBON 2	-	-2.2	-1.158	-	-
JEDKOT	s-*cis*(1)	-	$D_1^1(2)^*$	-	-175.763	7.501	-	-
KACNAD	s-*cis*(2)	-	$C_1^1(7)$ - RIBBON 2	-	0	0	-	-
KEXXUG	s-*cis*(2)	-	$C_1^1(7)$ - RIBBON 2	-	-0.219	1.934	-	-
KIZNIQ	s-*trans*(1)	-	$C_1^1(7)$ - CHAIN	-	-179.972	-164.512	-	-
MENBIR	s-*cis*(1)	-	$C_1^1(7)$ - RIBBON 1	-	179.98	5.786	-	-

MEQPED	s-*trans*(1)	-	$C_1^1(7)$ - CHAIN	-	177.554	-178.412	-	-
MIBYEB	s-*trans*(1)	-	$C_1^1(7)$ - CHAIN	-	172.314	-172.471	-	-
MPSIHF	s-*cis*(2)	-	$C_1^1(7)$ - RIBBON 2	-	5.087	13.286	-	-
NONRUD	s-*trans*(1)	-	$C_1^1(7)$ - RIBBON 1	-	-171.874	177.577	-	-
NUKWUL	s-*trans*(1)	-	$C_1^1(7)$ - RIBBON 1	-	175.867	-171.216	-	-
RABYID	s-*trans*(1)	-	$C_1^1(7)$ - RIBBON 1	-	179.781	179.408	-	-
RUXQIK	s-*cis*(1)	-	$C_1^1(7)$ - RIBBON 1	-	175.948	-3.652	-	-
TOYTOQ	s-*trans*(1)	-	$D_1^1(2)^*$	-	175.741	-179.782	-	-
WAVCEB	s-*cis*(1)	-	$C_1^1(7)$ - RIBBON 1	-	-163.065	-25.218	-	-
XINSAO	s-*cis*(1)	s-*cis*(1)	$C_1^1(7)$ - CHAIN	$C_1^1(7)$ - CHAIN	-172.929	-6.774	-176.884	-10.95
YIPSEV	s-*trans*(1)	-	$C_1^1(7)$ - RIBBON 1	-	-175.419	-168.963	-	-
GOLZIR	s-*cis*(1)	-	$C_1^1(7)$ - CHAIN	-	175.619	-9.091	-	-
SOLYIC	s-*cis*(2)	-	$C_1^1(7)$ - RIBBON 2	-	0	-1.203	-	-
POWMOE	s-*cis*(2)	-	$C_1^1(7)$ - RIBBON 2	-	-2.789	17.666	-	-

Table 7.4: Tabulated conformational types and torsion angles for the fumarate salts retrieved from the Cambridge Structural Database. * These structures do not have any fumarate-fumarate hydrogen bond interactions. Instead they are hydrated salt structures where the COOH group of the monoionised fumarate ion forms hydrogen bonds with a bridging water molecule. Structures with two crystallographically distinct fumarate ions in the asymmetric unit form hydrogen bonds such as the $C_2^2(14)$ motif of **AMAVEP** which are not observed in predicted 1:1 crystal energy landscape of 1,8-naphthyridinium fumarate, **XIX**.

8 Overall conclusions and suggestions for further work

Current industrial practice relies on choosing suitable acid-base pairs for salt formation based on the relative solution ionisation constants. In cases where such ionisation constants fail to predict salt formation, it is not uncommon for industrial salt selection programs to resort to trial and error, where previous experience plays an important role in narrowing the choice of *GRAS* counterions to screen. This thesis has applied computational crystal structure prediction to organic salts containing the chloride and carboxylate counterions and contributes to a larger inter-institutional project called *Control and Prediction of the Organic Solid State*. A method of crystal structure prediction based on the global minimisation of the lattice energy was tested against the known crystal structures of chloride and carboxylate salts. This thesis has examined the insights offered by computational crystal structure prediction when applied to two problems commonly encountered in salt crystallisation screens. These are the problems of solid state proton disorder and hydration of salt crystal structures. Finally an objective assessment of the success of the computational method used in this thesis was performed by participating in the fifth international blind test of crystal structure prediction and performing the predictions under blind conditions.

8.1 Success in predicting experimentally known salt structures

The CrystalPredictor/DMACRYS workflow of crystal structure prediction (CSP) was used to calculate the crystal energy landscapes of selected systems, keeping the molecules/ions rigid in the search (Chapters 5, 6 and 7) and relaxing any flexible torsion angles as an additional step with CrystalOptimizer (Chapter 7). The experimental structures of these model systems (some were cocrystals) were known prior to the modelling and this provided the basis for validating the computational method of CSP. For the majority of systems, this systematic global lattice energy optimisation search algorithm proved successful in locating the experimental structure as the most stable or energetically competitive with the most stable structure in the calculated crystal energy landscape. For example, the observed crystal structures of 4-dimethylaminopyridinium maleate (Chapter 5) and amantadine hydrochloride (Chapter 6) were predicted as the most stable structures in their respective crystal energy landscapes in agreement with

the results of manual crystallisation screens which suggested the observed structure as the only or most stable solid form respectively. For the disordered pyridine isophthalic acid system (Chapter 5), the cocrystal major product was predicted as energetically competitive with the global minimum and the crystal structure of 4-cyanopyridine 4-fluorobenzoic acid was predicted to be within 2 kJ mol^{-1} of the global lattice energy minimum. The work on chloride salts (Chapter 6) has shown that the CrystalPredictor/DMACRYS workflow of CSP can be applied to salts where one of the ions is a spherical point charge and future work would extend the calculations to the structures of metal containing salts. This will be demanding of accurate intermolecular potentials that may need to be developed as the work on chloride salts has shown the scarcity of intermolecular potentials for spherical ions that have been specifically fitted to the crystal structures of organic salts.

Given the underlying assumptions behind the model for the lattice energies, this is a good set of results to begin with and further improvements in the relative lattice energies could be obtained by estimating the induction contributions towards the lattice energy explicitly as this has now been implemented in DMACRYS. This would require the *ab initio* derivation of a model potential as the empirical potentials used in this thesis have absorbed the effects of induction from the fitting. Periodic density functional theory (DFT) calculations may also prove successful in accurately estimating the relative lattice energies of salts. This is because the dominant electrostatic contribution towards the lattice energies of salts is known to be modelled accurately by this method. The method requires a suitable dispersion-correction for the calculated lattice energies as DFT is known to treat dispersion interactions poorly. The high computational cost of DFT calculations will also be a limitation.

8.2 Solid state proton disorder and insights from computational crystal structure prediction

The difficulties in locating acidic protons from X-ray diffraction experiments and the failure of solution ionisation constants to correctly predict salt formation for some acid-base pairs provided the motivation to ask the question: Does it really matter where the acidic proton is in the crystal structures of salts and cocrystals? The answer is a definitive yes according to the results (Chapter 4) of two contrasting computational

modelling methods performed on a set of 9 crystal structures, 6 of which were discovered as part of extensive salt crystallisation screens on pyridine and 4-dimethylaminopyridine. For the majority of systems, there were significant structural differences between the rigid body lattice energy minima calculated with the salt and cocrystal molecular structures and better structural agreement with experiment was observed when the correct proton position was used. Although the rigid body modelling work proved successful for the particular systems tested, the results of such modelling are known to be sensitive to the intermolecular force field used. In selected systems, fixed cell periodic electronic structure calculations have proven to be a useful diagnostic tool in suggesting the energetic preference for a salt or cocrystal although this comes at a considerable computational cost even for reasonably sized molecules/ions.

The research presented in Chapter 5 has shown that the results of CSP are sensitive to the atomic connectivity of the acidic proton in the crystal. This was evident in the re-ranking of structures based on the same hydrogen bond motif and the finding of new motifs when separate predictions were performed using the molecular structures of the salt and cocrystal. For all systems, the observed crystal structure had a more favourable lattice energy relative to the predicted global minimum when compared to the results for the hypothetical structure. This remains a qualitative assessment as the isolated molecule/ion conformations used in the predictions mean that an estimate of the true energy penalty for proton transfer is outside the scope of the CrystalPredictor/ DMACRYS suite of programs. Despite this limitation, the sensitivity of the calculated salt and cocrystal lattice energy minima to the atomic connectivity of the acidic proton, meant that comparisons of the root-mean-square deviations of these structures from the reference experimental structure was diagnostic enough in the majority of cases. The caveat that goes with this sort of analysis is that it is demanding of a sufficiently good model for the intermolecular forces of the salt and cocrystal.

8.3 Using crystal structure prediction to rationalise the hydration behaviour of organic salts

Hydration of chloride salts is a commonly encountered phenomenon and this thesis has shown that it is possible to infer the likelihood of crystallising the salt as a hydrate by

looking at the solvent accessible voids in the predicted low energy structures of the anhydrous salt. For memantine hydrochloride, the fact that all predicted structures within 12 kJ mol^{-1} of the global minimum had solvent accessible voids was taken to be a warning that the salt is susceptible to hydration and this was later confirmed by the results of manual crystallisation screens. For amantadine hydrochloride, only two predicted structures within 12 kJ mol^{-1} of the global minimum had solvent accessible voids and the experimental structure was not among them. This suggested minimal risk of hydration and was confirmed by the results of experimental screens.

The accuracy of the assumed molecular structures is just as important as the model for the lattice energies, in affecting the results of CSP. Indeed, for memantine hydrochloride, the fact that the experimental structure is predicted to be 7.33 kJ mol^{-1} higher in energy than the global minimum, could be explained if the channels in the structure are occupied by disordered water molecules as suggested by the residual electron density from the crystallography.

8.4 The fifth international blind test of crystal structure prediction

The majority of the predictions in this thesis were performed using known crystal structures and this was necessary for the successful validation of the computational method. The fifth international blind test of crystal structure provided an ideal setting to test the computational method under truly blind conditions. This was the first blind test where the molecular structure of a salt, 1,8-naphthyridinium fumarate (**XIX**), was proposed as a target for researchers. The fact that the experimental structure of **XIX** was not predicted in the search was due to an insufficient number of CrystalPredictor minimisations. The finding that the 1,8-naphthyridinium fumarate salt, **XIX**, and its isostructural counterpart quinolinium fumarate, **RABYID**, have unfavourable lattice energies in their respective crystal energy landscapes was not unique to the approach used in this thesis. In fact, a workshop organised by the Cambridge Crystallographic Data Centre revealed that the groups of Kendrick and Day also observed similar results for these salts and the majority of participants did not have the experimental structure in the extended list of predicted lattice energy minima. Only van Eijck was successful in predicting the experimental structure of **XIX** as one of the three lowest energy structures and Kendrick included the experimental structure among his submissions on

the basis of its isostructurality to **RABYID**. The success of van Eijck is likely due to his use of the dimer wavefunction to derive the atomic multipole moments for the ions. This differs from the approach used in this thesis as it includes the polarisation of the ions by the hydrogen bonding within the dimer.

In this thesis, the majority of the successes in CSP have been for salts with one dominant hydrogen bond between the oppositely charged ions. However, the lesson from the work on 1,8-naphthyridinium fumarate, **XIX**, is that when there are many competing intermolecular interactions in the crystal, empirically derived dispersion-repulsion potentials are not reliable for estimating the relative energies of predicted structures. Thus future work on salts with mono-ionised dicarboxylic acids may benefit from the use of *ab initio* derived model potentials and an explicit treatment of induction effects.

8.5 Summary

Overall, this thesis has shown that CSP can serve as a complementary tool to experimental efforts in salt screening. The agreement between theory and experiment obtained following the modelling of selected salt structures suggests that CSP could be used to rationalise the difficulties in crystallising organic salts. However future work should look to improve the estimates in the relative energies of predicted structures. This may be done by deriving the intermolecular potentials *ab initio* and combining this with a suitable model for the induction and electrostatic contributions or by moving towards dispersion-corrected density functional theory methods.

Bibliography

Reference List

1. Gardner, C. R.; Walsh, C. T.; Almarsson, O. Drugs as materials: valuing physical form in drug discovery. *Nat Rev Drug Discov* **2004,** *3* (11), 926-934.

2. Agharkar, S.; Lindenbaum, S.; Higuchi, T. Enhancement of Solubility of Drug Salts by Hydrophilic Counterions - Properties of Organic Salts of An Antimalarial Drug. *J. Pharm. Sci.* **1976,** *65* (5), 747-749.

3. Aaltonen, J.; Alleso, M.; Mirza, S.; Koradia, V.; Gordon, K. C.; Rantanen, J. Solid form screening - A review. *Eur. J. Pharm. Biopharm.* **2009,** *71* (1), 23-37.

4. *Polymorphism in the Pharmaceutical Industry;* Wiley-VCH: Weinheim, 2006.

5. Remenar, J. F.; Morissette, S. L.; Peterson, M. L.; Moulton, B.; MacPhee, J. M.; Guzman, H. R.; Almarsson, O. Crystal engineering of novel cocrystals of a Triazole drug with 1,4-Dicarboxylic acids. *J. Am. Chem. Soc.* **2003,** *125*, 8456-8457.

6. Good, D. J.; Rodriguez-Hornedo, N. Solubility Advantage of Pharmaceutical Cocrystals. *Cryst. Growth Des.* **2009,** *9* (5), 2252-2264.

7. Berry, D. J.; Seaton, C. C.; Clegg, W.; Harrington, R. W.; Coles, S. J.; Horton, P. N.; Hursthouse, M. B.; Storey, R.; Jones, W.; Friscic, T.; Blagden, N. Applying Hot-Stage Microscopy to Co-Crystal Screening: A Study of Nicotinamide with Seven Active Pharmaceutical Ingredients. *Cryst. Growth Des.* **2008,** *8* (5), 1697-1712.

8. Batchelor, E.; Klinowski, J.; Jones, W. Crystal engineering using co-crystallisation of phenazine with dicarboxylic acids. *J. Mater. Chem.* **2000,** *10* (4), 839-848.

9. Desiraju, G. R. Crystal engineering. From molecules materials. *Journal of Molecular Structure* **2003,** *656* (1-3), 5-15.

10. Sarma, J. A. R. P.; Desiraju, G. R. The supramolecular synthon approach to crystal structure prediction. *Cryst. Growth Des.* **2002,** *2* (2), 93-100.

11. Desiraju, G. R. Chemistry beyond the molecule. *Nature* **2001,** *412* (6845), 397-400.

12. Dunitz, J. D. Crystal and co-crystal: a second opinion. *CrystEngComm* **2003,** *5*, 506.

13. Desiraju, G. R. Crystal and co-crystal. *CrystEngComm* **2003,** *5* (82), 466-467.

14. Aakeroy, C. B.; Salmon, D. J. Building co-crystals with molecular sense and supramolecular sensibility. *CrystEngComm* **2005,** *7* (72), 439-448.

15. Bond, A. D. What is a co-crystal? *CrystEngComm* **2007,** *9* (9), 833-834.

16. Wiechert, D.; Mootz, D. Molecular beside ionic: Crystal structures of a 1/1 and a 1/4 adduct of pyridine and formic acid. *Angew. Chem. ,Int. Ed.* **1999,** *38* (13-14), 1974-1976.

17. Shan, N.; Zaworotko, M. J. The role of cocrystals in pharmaceutical science. *Drug Discover. Today* **2008,** *13* (9-10), 440-446.

18. Steiner, T. The hydrogen bond in the solid state. *Angew. Chem. ,Int. Ed.* **2002,** *41* (1), 48-76.

19. Pauling, L. The structure and entropy of ice and other crystals with some randomness of atomic arrangement. *J. Am. Chem. Soc.* **1935,** *57*, 2680-2684.

20. Pauling, L. *The Nature of the Chemical Bond;* Third ed.; Cornell University Press: New York, 1960.

21. Desiraju, G. R.; Steiner, Th. *The Weak Hydrogen Bond.;* Oxford University Press: Oxford, 1999.

22. Etter, M. C. Encoding and decoding hydrogen-bond patterns of organic compounds. *Accounts Chem. Res.* **1990,** *23* (4), 120-126.

23. Aakeroy, C. B.; Beatty, A. M.; Helfrich, B. A. A high-yielding supramolecular reaction. *J. Am. Chem. Soc.* **2002,** *124* (48), 14425-14432.

24. McMahon, J. A.; Bis, J. A.; Vishweshwar, P.; Shattock, T. R.; McLaughlin, O. L.; Zaworotko, M. Crystal engineering of the composition of pharmaceutical phases. 3.Primary amide supramolecular heterosynthons and their role in the design of pharmaceutical co-crystal. *Z. Kristallogr.* **2005,** *220*, 340.

25. Bernstein, J.; Davies, R. E.; Shimoni, L.; Chang, N. Patterns in Hydrogen Bonding: Functionality and graph set analysis in crystals. *Angew. Chem. ,Int. Ed. Engl.* **1995,** *34* (15), 1555-1573.

26. Desiraju, G. R. Hydrogen bridges in crystal engineering: Interactions without borders. *Accounts Chem. Res.* **2002,** *35* (7), 565-573.

27. Bhogala, B. R.; Nangia, A. Cocrystals of 1,3,5-cyclohexanetricarboxylic acid with 4,4 '-bipyridine homologues: Acid center dot center dot center dot pyridine hydrogen bonding in neutral and ionic complexes. *Cryst. Growth Des.* **2003,** *3* (4), 547-554.

28. Bis, J. A.; Zaworotko, M. J. The 2-aminopyridinium-carboxylate supramolecular heterosynthon: A robust motif for generation of multiple-component crystals. *Cryst. Growth Des.* **2005,** *5* (3), 1169-1179.

29. Allen, F. H.; Taylor, R. Research applications of the Cambridge Structural Database (CSD). *Chem. Soc. Rev.* **2004,** *33* (8), 463-475.

30. Etter, M. C.; MacDonald, J. C.; Bernstein, J. Graph-Set Analysis of Hydrogen-Bond Patterns in Organic- Crystals. *Acta Crystallogr. , Sect. B* **1990,** *46*, 256-262.

31. Bernstein, J. *Polymorphism in Molecular Crystals;* Clarendon Press: Oxford, 2002.

32. E.Mitscherlich *Abhl. Akad. Berlin* **1823,** *43*.

33. Chemburkar, S. R.; Bauer, J.; Deming, K.; Spiwek, H.; Patel, K.; Morris, J.; Henry, R.; Spanton, S.; Dziki, W.; Porter, W.; Quick, J.; Bauer, P.; Donaubauer, J.; Narayanan, B. A.; Soldani, M.; Riley, D.; McFarland, K. Dealing with the impact of ritonavir polymorphs on the late stages of bulk drug process development. *Org. Process Res. Dev.* **2000,** *4* (5), 413-417.

34. Gavezzotti, A. Calculation of intermolecular interaction energies by direct numerical integration over electron densities. I. Electrostatic and polarization energies in molecular crystals. *J. Phys. Chem. B* **2002,** *106* (16), 4145-4154.

35. Gavezzotti, A. Ten years of experience in polymorph prediction: what next? *CrystEngComm* **2002,** *4* (61), 343-347.

36. Burger, M. J.; Bloom, M. C. *Z. Kristallogr.* **1937,** *96*.

37. Einstein, A. *Ann. Phys.* **1907,** *22*, 180-190.

38. Ostwald, W. Z. Studies on the formation and transformation of solid phases. *Z. Phys. Chem.* **1897,** *22*, 289-330.

39. Miers, H. A.; Isaac, F. The spontaneous crystallisation of binary mixtures. *Proceedings of the Royal Society* **1907,** *A79*, 322-351.

40. Mullin, J. W. *Crystallization;* 3rd ed.; Butterworh-Heinmann: Oxford, 1993.

41. Dunitz, J. D.; Bernstein, J. Disappearing Polymorphs. *Accounts Chem. Res.* **1995,** *28* (4), 193-200.

42. Lancaster, R. W.; Karamertzanis, P. G.; Hulme, A. T.; Tocher, D. A.; Lewis, T. C.; Price, S. L. The Polymorphism of Progesterone: Stabilization of a 'Disappearing' Polymorph by Co-Crystallization. *J. Pharm. Sci.* **2007,** *96* (12), 3419-3431.

43. Blagden, N.; Davey, R. J.; Rowe, R.; Roberts, R. Disappearing polymorphs and the role of reaction by-products: the case of sulphathiazole. *Int. J. Pharm.* **1998,** *172* (1-2), 169-177.

44. Hulliger, J. Chemistry and Crystal-Growth. *Angew. Chem. ,Int. Ed. Engl.* **1994,** *33* (2), 143-162.

45. Clegg, W.; Blake, A. J.; Gould, R. O.; Main, P. *Crystal Structure Analysis, Principles and Practice;* Oxford Science Publications: Oxford, 2001.

46. Friscic, T. New opportunities for materials synthesis using mechanochemistry. *J. Mater. Chem.* **2010,** *20* (36), 7599-7605.

47. Trask, A. V.; van de Streek, J.; Motherwell, W. D. S.; Jones, W. Achieving polymorphic and stoichiometric diversity in cocrystal formation: Importance of solid-state grinding, powder X-ray structure determination, and seeding. *Cryst. Growth Des.* **2005,** *5* (6), 2233-2241.

48. Trask, A. V.; Jones, W. Crystal engineering of organic cocrystals by the solid-state grinding approach. *Top. Curr. Chem.* **2005,** *254*, 41-70.

49. Patil, A. O.; Curtin, D. Y.; Paul, I. C. Solid-State Formation of Quinhydrones from Their Components - Use of Solid-Solid Reactions to Prepare Compounds Not Accessible from Solution. *J. Am. Chem. Soc.* **1984,** *106* (2), 348-353.

50. Day, G. M.; Trask, A. V.; Motherwell, W. D. S.; Jones, W. Investigating the latent polymorphism of maleic acid. *Chem. Commun.* **2006,** (1), 54-56.

51. Rafilovich, M.; Bernstein, J. Serendipity and four polymorphic structures of benzidine, C12H12N2. *Journal of the American Chemical Society* **2006,** *128* (37), 12185-12191.

52. Braga, D.; Grepioni, F. Making crystals from crystals: a green route to crystal engineering and polymorphism. *Chem. Commun.* **2005,** (29), 3635-3645.

53. Trask, A. V.; Haynes, D. A.; Motherwell, W. D. S.; Jones, W. Screening for crystalline salts via mechanochemistry. *Chem. Commun.* **2006,** (1), 51-53.

54. Friscic, T.; Jones, W. Recent Advances in Understanding the Mechanism of Cocrystal Formation via Grinding. *Cryst. Growth Des.* **2009,** *9* (3), 1621-1637.

55. Cruz-Cabeza, A. J.; Karki, S.; Fabian, L.; Friscic, T.; Day, G. M.; Jones, W. Predicting stoichiometry and structure of solvates. *Chem. Commun.* **2010,** *46* (13), 2224-2226.

56. Di Profio, G.; Curcio, E.; Drioli, E. A Review on membrane crystallization. *Chim. Oggi* **2009,** *27* (2), 27-31.

57. Curcio, E.; Fontananova, E.; Di Profio, G.; Drioli, E. Influence of the structural properties of poly(vinylidene fluoride) membranes on the heterogeneous nucleation rate of protein crystals. *Journal of Physical Chemistry B* **2006,** *110* (25), 12438-12445.

58. Di Profio, G.; Tucci, S.; Curcio, E.; Drioli, E. Selective glycine polymorph crystallization by using microporous membranes. *Crystal Growth & Design* **2007,** *7* (3), 526-530.

59. Bucar, D. K.; MacGillivray, L. R. Preparation and reactivity of nanocrystalline cocrystals formed via sonocrystallization. *Journal of the American Chemical Society* **2007,** *129* (1), 32-33.

60. Stuart, M.; Box, K. Chasing equilibrium: Measuring the intrinsic solubility of weak acids and bases. *Anal. Chem.* **2005,** *77* (4), 983-990.

61. Avdeef, A.; Box, K. J.; Comer, J. E. A.; Hibbert, C.; Tam, K. Y. pH-metric logP 10. Determination of liposomal membrane-water partition coefficients of ionizable drugs. *Pharmaceutical Research* **1998,** *15* (2), 209-215.

62. Llinas, A.; Box, K. J.; Burley, J. C.; Glen, R. C.; Goodman, J. M. A new method for the reproducible generation of polymorphs: two forms of sulindac with very different solubilities. *Journal of Applied Crystallography* **2007,** *40*, 379-381.

63. Blundell, T. L.; Jhoti, H.; Abell, C. High-throughput crystallography for lead discovery in drug design. *Nature Reviews Drug Discovery* **2002,** *1* (1), 45-54.

64. Stahl, P. H.; Wermuth (Eds.), C. G. *Handbook of Pharmaceutical Salts: Properties, Selection, and Use;* Verlag Helvetica Chimica Acta, Zurich/Wiley-VCH: 2002.

65. Remenar, J. F.; MacPhee, J. M.; Larson, B. K.; Tyagi, V. A.; Ho, J. H.; McIlroy, D. A.; Hickey, M. B.; Shaw, P. B.; Almarsson, O. Salt selection and simultaneous polymorphism assessment via high-throughput crystallization: The case of sertraline. *Organic Process Research & Development* **2003,** *7* (6), 990-996.

66. Bowker, M. J. A procedure for salt selection and optimization. In *Handbook of Pharmaceutical Salts*, Stahl, P. H., Wermuth, C. G., Eds.; VHCA and Wiley-VCH: 2002.

67. Smith, R. L.; Cohen, S. M.; Doull, J.; Feron, V. J.; Goodman, J. I.; Marnett, L. J.; Portoghese, P. S.; Waddell, W. J.; Wagner, B. M.; Adams, T. B. GRAS flavoring substances 21. *Food Technology* **2003,** *57* (5), 46.

68. Verbeeck, R. K.; Kanfer, I.; Walker, R. B. Generic substitution: The use of medicinal products containing different salts and implications for safety and efficacy. *European Journal of Pharmaceutical Sciences* **2006,** *28* (1-2), 1-6.

69. Serajuddin, A. T. M.; Puddipeddi, M. Salt selection strategies. In *Handbook of Pharmaceutical Salts*, Stahl, P. H., Wermuth, C. G., Eds.; VHCA and Wiley-VCH: Weinheim, 2002.

70. Bhogala, B. R.; Basavoju, S.; Nangia, A. Tape and layer structures in cocrystals of some di- and tricarboxylic acids with 4,4 '-bipyridines and isonicotinamide. From binary to ternary cocrystals. *CrystEngComm* **2005,** *7*, 551-562.

71. Tong, W. Q.; Whitesell, G. In situ salt screening - A useful technique for discovery support and preformulation studies. *Pharmaceutical Development and Technology* **1998,** *3* (2), 215-223.

72. He, G. W.; Chow, P. S.; Tan, R. B. H. Predicting Multicomponent Crystal Formation: The Interplay between Homomeric and Heteromeric Interactions. *Cryst. Growth Des.* **2009,** *9* (10), 4529-4532.

73. Haynes, D. A.; Jones, W.; Motherwell, W. D. S. Hydrate formation in NH+ -containing salts of pharmaceutically acceptable anions: A CSD survey. *CrystEngComm* **2005,** *7* (55), 342-345.

74. Barton, J. H. Changing intellectual property issues in the biotechnology industry. *Biotechnology Law Report* **1999,** *18* (1), 3-13.

75. Price, S. L. From crystal structure prediction to polymorph prediction: interpreting the crystal energy landscape. *Phys. Chem. Chem. Phys.* **2008,** *10* (15), 1996-2009.

76. Filippini, G.; Gavezzotti, A. Empirical Intermolecular Potentials For Organic-Crystals: the '6-Exp' Approximation Revisited. *Acta Crystallogr. , Sect. B* **1993,** *49* (Pt5), 868-880.

77. Hulme, A. T.; Price, S. L.; Tocher, D. A. A New Polymorph of 5-Fluorouracil Found Following Computational Crystal Structure Predictions. *J. Am. Chem. Soc.* **2005,** *127* (4), 1116-1117.

78. Beyer, T.; Day, G. M.; Price, S. L. The prediction, morphology, and mechanical properties of the polymorphs of paracetamol. *J. Am. Chem. Soc.* **2001,** *123* (21), 5086-5094.

79. Cruz-Cabeza, A. J.; Day, G. M.; Jones, W. Towards Prediction of Stoichiometry in Crystalline Multicomponent Complexes. *Chem. Eur. J* **2008,** *14* (29), 8830-8836.

80. Hulme, A. T.; Price, S. L. Towards the prediction of organic hydrate crystal structures. *J. Chem. Theory Comput.* **2007,** *3* (4), 1597-1608.

81. Hulme, A. T.; Tocher, D. A. The Discovery of New Crystal Forms of 5-Fluorocytosine Consistent with the Results of Computational Crystal Structure Prediction. *Cryst. Growth Des.* **2006,** *6* (2), 481-487.

82. Schmidt, M. U. Energy minimisation as a tool for crystal structure determination of industrial pigments. In *Crystal Engineering: From Molecules and Crystals to Materials*, Braga, D., Grepioni, F., Orpen, A. G., Eds.; Kluwer: Dordrecht, 1999; pp 331-348.

83. Laio, A.; Parrinello, M. Escaping free-energy minima. *P. Natl. Acad. Sci. USA* **2002,** *99* (20), 12562-12566.

84. Raiteri, P.; Martonak, R.; Parrinello, M. Exploring Polymorphism: The Case of Benzene. *Angew. Chem. ,Int. Ed.* **2005,** *44*, 3769-3773.

85. van Eijck, B. P.; Spek, A. L.; Mooij, W. T. M.; Kroon, J. Hypothetical crystal structures of benzene at 0 and 30 kbar. *Acta Crystallogr. , Sect. B* **1998,** *54* (3), 291-299.

86. Karamertzanis, P. G.; Raiteri, P.; Parrinello, M.; Leslie, M.; Price, S. L. The Thermal Stability of Lattice Energy Minima of 5-Fluorouracil: Metadynamics as an Aid to Polymorph Prediction. *J. Phys. Chem. B* **2008,** *112* (14), 4298-4308.

87. Stone, A. J. *The Theory of Intermolecular Forces;* 1st ed.; Clarendon Press: Oxford, 1996.

88. Karamertzanis, P. G.; Pantelides, C. C. Ab initio crystal structure prediction - I. Rigid molecules. *J. Comput. Chem.* **2005,** *26* (3), 304-324.

89. Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. Modelling organic crystal structures using distributed multipole and polarizability-based model intermolecular potentials. *Phys. Chem. Chem. Phys.* **2010,** *12* (30), 8478-8490.

90. Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor. Chim. Acta* **1977,** *44* (2), 129-138.

91. Nalewajski, R. F.; Broniatowska, E. Atoms-in-molecules from the stockholder partition of the molecular two-electron distribution. *Theoretical Chemistry Accounts* **2007,** *117* (1), 7-27.

92. Bader, R. F. W. *Atoms in Molecules. A Quantum Theory;* Clarendon Press: Oxford, 1990; Vol. 22.

93. Breneman, C. M.; Wiberg, K. B. Determining Atom-Centered Monopoles From Molecular Electrostatic Potentials - the Need For High Sampling Density in Formamide Conformational-Analysis. *J. Comput. Chem.* **1990,** *11* (3), 361-373.

94. Stone, A. J. Distributed Multipole Analysis, or How to Describe a Molecular Charge Distribution. *Chem. Phys. Lett.* **1981,** *83* (2), 233-239.

95. Stone, A. J.; Alderton, M. Distributed Multipole Analysis - Methods and Applications. *Mol. Phys.* **1985,** *56* (5), 1047-1064.

96. Day, G. M.; Motherwell, W. D. S.; Jones, W. Beyond the isotropic atom model in crystal structure prediction of rigid molecules: Atomic multipoles versus point charges. *Cryst. Growth Des.* **2005,** *5* (3), 1023-1033.

97. Oddershede, J.; Larsen, S. Charge density study of naphthalene based on X-ray diffraction data at four different temperatures and theoretical calculations. *J. Phys. Chem. A* **2004,** *108* (6), 1057-1063.

98. Welch, G. W. A.; Karamertzanis, P. G.; Misquitta, A. J.; Stone, A. J.; Price, S. L. Is the induction energy important for modeling organic crystals? *J. Chem. Theory Comput.* **2008,** *4* (3), 522-532.

99. London, F. *Trans. Faraday Soc.* **1937,** *33*, 4.

100. Williams, D. E.; Starr, T. L. Calculation of the Crystal Structures of Hydrocarbons by Molecular Packing Analysis. *Computers & Chemistry* **1977,** *1*, 173-177.

101. Hsu, L. Y.; Williams, D. E. Intermolecular Potential-Function Models for Crystalline Perchlorohydrocarbons. *Acta Crystallogr. , Sect A.* **1980,** *36* (MAR), 277-281.

102. Cox, S. R.; Hsu, L. Y.; Williams, D. E. Nonbonded Potential Function Models for Crystalline Oxohydrocarbons. *Acta Crystallogr. , Sect A.* **1981,** *37* (MAY), 293-301.

103. Williams, D. E.; Cox, S. R. Nonbonded Potentials For Azahydrocarbons: the Importance of the Coulombic Interaction. *Acta Crystallogr. , Sect. B* **1984,** *40* (8), 404-417.

104. Williams, D. E.; Houpt, D. J. Fluorine Nonbonded Potential Parameters Derived From Crystalline Perfluorocarbons. *Acta Crystallogr. , Sect. B* **1986,** *42* (JUN), 286-295.

105. Coombes, D. S.; Price, S. L.; Willock, D. J.; Leslie, M. Role of Electrostatic Interactions in Determining the Crystal Structures of Polar Organic Molecules. A Distributed Multipole Study. *J. Phys. Chem.* **1996,** *100* (18), 7352-7360.

106. Beyer, T.; Price, S. L. Dimer or catemer? Low-energy crystal packings for small carboxylic acids. *J. Phys. Chem. B* **2000,** *104* (12), 2647-2655.

107. Williams, D. E. Improved intermolecular force field for molecules containing H, C, N, and O atoms, with application to nucleoside and peptide crystals. *J. Comput. Chem.* **2001,** *22* (11), 1154-1166.

108. Beyer, T.; Price, S. L. The errors in lattice energy minimisation studies: sensitivity to experimental variations in the molecular structure of paracetamol. *CrystEngComm* **2000,** *2* (34), 183-190.

109. Karfunkel, H. R.; Gdanitz, R. J. Abinitio Prediction of Possible Crystal-Structures for General Organic-Molecules. *J. Comput. Chem.* **1992,** *13* (10), 1171-1183.

110. Holden, J. R.; Du, Z. Y.; Ammon, H. L. Prediction of Possible Crystal-Structures For C-, H- , N- , O- and F-Containing Organic- Compounds. *Journal of Computational Chemistry* **1993,** *14* (4), 422-437.

111. Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. , Sect. B* **2002,** *58* (3), 380-388.

112. Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical methods: the art of scientific computing (Fortran version);* Cambridge University Press: Cambridge, 1989.

113. Lecuyer, P. Efficient and Portable Combined Random Number Generators. *Communications of the Acm* **1988,** *31* (6), 742-&.

114. Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. Mercury CSD 2.0 - new features for the visualization and investigation of crystal structures. *J. Appl. Crystallogr.* **2008,** *41*, 466-470.

115. Chisholm, J. A.; Motherwell, S. COMPACK: a program for identifying crystal structure similarity using distances. *J. Appl. Crystallogr.* **2005,** *38*, 228-231.

116. Kazantsev, A. V.; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C. CrystalOptimizer. An efficient Algorithm for Lattice Energy Minimisation of Organic Crystal using Isolated-Molecule Quantum Mechanical Calculations. In *Molecular System Engineering*, Adjiman, C. S., Galindo, A., Eds.; WILEY-VCH Verlag GmbH & Co.: Weinheim, 2010; pp 1-42.

117. Leslie, M. A Symmetry-Adapted Method for the Determination of the Lattice Energy and Properties of Ionic-Crystals. *Solid State Ionics* **1983,** *8* (3), 243-246.

118. Smith, W. Point multipoles in the Ewald summation. *CCP5 Quarterly* **1982,** *4*, 13-25.

119. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Gaussian Inc.: Wallingford CT, **2003**

120. Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular Mechanics - the Mm3 Force-Field for Hydrocarbons .1. *J. Am. Chem. Soc.* **1989,** *111* (23), 8551-8566.

121. Karamertzanis, P. G.; Price, S. L. Energy Minimization of Crystal Structures Containing Flexible Molecules. *J. Chem. Theory Comput.* **2006,** *2* (4), 1184-1199.

122. Karamertzanis, P. G.; Kazantsev, A. V.; Issa, N.; Welch, G. W. A.; Adjiman, C. S.; Pantelides, C. C.; Price, S. L. Can the Formation of Pharmaceutical Co-Crystals Be Computationally Predicted? II. Crystal Structure Prediction. *J. Chem. Theory Comput.* **2009,** *5* (5), 1432-1448.

123. Day, G. M. Lattice dynamical studies of molecular crystals with application to polymorphism and structure prediction. Ph.D., University College London, 2003.

124. Polito, M.; D'Oria, E.; Maini, L.; Karamertzanis, P. G.; Grepioni, F.; Braga, D.; Price, S. L. The crystal structures of chloro and methyl ortho-benzoic acids and their co-crystal: rationalizing similarities and differences. *CrystEngComm* **2008,** *10*, 1848-1854.

125. Habgood, M.; Deij, M. A.; Mazurek, J.; Price, S. L.; ter Horst, J. H. Carbamazepine Co-crystallization with Pyridine Carboxamides: Rationalization by Complementary Phase Diagrams and Crystal Energy Landscapes. *Cryst. Growth Des.* **2009,** *10* (2), 903-912.

126. Habgood, M.; Price, S. L. Isomers, Conformers, and Cocrystal Stoichiometry: Insights from the Crystal Energy Landscapes of Caffeine with the Hydroxybenzoic Acids. *Cryst. Growth Des.* **2010,** *10* (7), 3263-3272.

127. Aakeroy, C. B.; Desper, J.; Fasulo, M. E. Improving success rate of hydrogen-bond driven synthesis of co-crystals. *CrystEngComm* **2006,** *8* (8), 586-588.

128. Aakeroy, C. B.; Hussain, I.; Forbes, S.; Desper, J. Exploring the hydrogen-bond preference of N-H moieties in co-crystals assembled via O-H(acid)N(py) intermolecular interactions. *CrystEngComm* **2007,** *9* (1), 46-54.

129. Vishweshwar, P.; McMahon, J. A.; Bis, J. A.; Zaworotko, M. J. Pharmaceutical co-crystals. *J. Pharm. Sci.* **2006,** *95* (3), 499-516.

130. Wishkerman, S.; Bernstein, J.; Hickey, M. B. Crystal Engineering with Cocrystals of Benzo-[18]Crown-6 and Urea and Thiourea Derivatives. *Cryst. Growth Des.* **2009**.

131. Karamertzanis, P. G.; Anandamanoharan, P. R.; Fernandes, P.; Cains, P. W.; Vickers, M.; Tocher, D. A.; Florence, A. J.; Price, S. L. Toward the computational design of diastereomeric resolving agents: An experimental and computational study of 1-phenylethylammonium-2-phenylacetate derivatives. *J. Phys. Chem. B* **2007,** *111* (19), 5326-5336.

132. Willock, D. J.; Price, S. L.; Leslie, M.; Catlow, C. R. A. The Relaxation of Molecular Crystal Structures Using a Distributed Multipole Electrostatic Model. *J. Comput. Chem.* **1995,** *16* (5), 628-647.

133. Antoniadis, C. D.; D'Oria, E.; Karamertzanis, P. G.; Tocher, D. A.; Florence, A. J.; Price, S. L.; Jones, A. G. A Computationally Inspired Investigation of the Solid Forms of (R)-1-Phenylethylammonium-(S)-2-phenylbutyrate. *Chirality* **2010,** *22* (4), 447-455.

134. Leusen, F. J. J. Crystal structure prediction of diastereomeric salts: A step toward rationalization of racemate resolution. *Cryst. Growth Des.* **2003,** *3* (2), 189-192.

135. McArdle, P.; Gilligan, K.; Cunningham, D.; Dark, R.; Mahon, M. A method for the prediction of the crystal structure of ionic organic compounds - the crystal structures of o-toluidinium chloride and bromide and polymorphism of bicifadine hydrochloride. *CrystEngComm* **2004,** *6* (53), 303-309.

136. Busing, W. R. *WMIN A Computer Program to Model Molecules and Crystals in Terms of Potential Energy Functions,* Oak Ridge National laboratory: Tennessee, USA, **1981**

137. Scott, R. A.; Scheraga, H. A. Method for Calculating Internal Rotation Barriers. *J. Chem. Phys.* **1965,** *42* (6), 2209-2215.

138. van de Streek, J.; Neumann, M. A.; Perrin, M. A. Validation of dispersion-corrected density functional theory calculations for the crystal structure prediction of molecular salts: a crystal structure prediction study of pyridinium chloride. *CrystEngComm* **2010**.

139. Neumann, M. A.; Perrin, M. A. Energy ranking of molecular crystals using density functional theory calculations and an empirical van der Waals correction. *J. Phys. Chem. B* **2005,** *109* (32), 15531-15541.

140. Neumann, M. A. *GRACE*, version 1.0; Avant-garde Materials Simulation: **2007**

141. Mootz, D.; Hocken, J. The System Pyridine-Hydrogen Chloride - Formation and Structure of Crystalline Adducts. *Z. Naturforsch. B* **1989,** *44* (10), 1239-1246.

142. Rerat, C. Structure Cristalline du Chlorhydrate de Pyridine. *Acta Crystallogr.* **1962,** *15* (MAY), 427.

143. Hensen, K.; Pullmann, P.; Bats, J. W. Crystal and Molecular-Structures of Chlorotrispyridinium-Bis-(Tetrachloroaluminate(Iii)) and A New Modification of Pyridinium Chloride. *Z. Anorg. Allg. Chem.* **1988,** *556* (1), 62-69.

144. Langkilde, A.; Oddershede, J.; Larsen, S. Diastereomeric salts of lactic acid and 1-phenylethylamine, their structures and relative stabilities. *Acta Crystallographica Section B-Structural Science* **2002,** *58*, 1044-1050.

145. Copyright CCDC 2001-2007 *Mercury CSD 2.0 (Build 4)*, **2008**

146. Pertsin, A. J.; Kitaigorodsky, A. I. *The Atom-Atom Potential Method. Applications to Organic Molecular Solids;* Springer-Verlag: Berlin, 1987; Vol. 43.

147. Williams, D. E. Improved intermolecular force field for crystalline hydrocarbons containing four- or three-coordinated carbon. *J. Mol. Struct.* **1999,** *486*, 321-347.

148. Spek, A. L. Single-crystal structure validation with the program PLATON. *J. Appl. Crystallogr.* **2003,** *36*, 7-13.

149. Hall, S. R.; Allen, F. H.; Brown, I. D. The Crystallographic Information File (Cif) - A New Standard Archive File for Crystallography. *Acta Cryst. A* **1991,** *47*, 655-685.

150. van de Streek, J. Searching the Cambridge Structural Database for the 'best' representative of each unique polymorph. *Acta Crystallogr. , Sect. B* **2006,** *62*, 567-579.

151. Harding, M. M. The geometry of metal-ligand interactions relevant to proteins. *Acta Crystallographica Section D-Biological Crystallography* **1999,** *55*, 1432-1443.

152. Haynes, D. A.; Jones, W.; Motherwell, W. D. S. Occurrence of pharmaceutically acceptable anions and cations in the Cambridge structural database. *J. Pharm. Sci.* **2005,** *94* (10), 2111-2120.

153. Desiraju, G. R. *Crystal Engineering: the Design of Organic Solids;* Elsevier: Amsterdam, 1989.

154. Blagden, N.; Berry, D. J.; Parkin, A.; Javed, H.; Ibrahim, A.; Gavan, P. T.; De Matos, L. L.; Seaton, C. C. Current directions in co-crystal growth. *New J. Chem.* **2008,** *32* (10), 1659-1672.

155. Shattock, T. R.; Arora, K. K.; Vishweshwar, P.; Zaworotko, M. J. Hierarchy of Supramolecular Synthons: Persistent Carboxylic Acid center dot center dot center dot Pyridine Hydrogen Bonds in Cocrystals That also Contain a Hydroxyl Moiety. *Cryst. Growth Des.* **2008,** *8* (12), 4533-4545.

156. Sarma, B.; Nath, N. K.; Bhogala, B. R.; Nangia, A. Synthon Competition and Cooperation in Molecular Salts of Hydroxybenzoic Acids and Aminopyridines. *Cryst. Growth Des.* **2009,** *9* (3), 1546-1557.

157. Anderson, K. M.; Probert, M. R.; Whiteley, C. N.; Rowland, A. M.; Goeta, A. E.; Steed, J. W. Designing Co-Crystals of Pharmaceutically Relevant Compounds That Crystallize with Z ' > 1. *Cryst. Growth Des.* **2009,** *9* (2), 1082-1087.

158. Aakeroy, C. B.; Hussain, I.; Forbes, S.; Desper, J. Exploring the hydrogen-bond preference of N-H moieties in co-crystals assembled via O-H(acid)center dot center dot center dot N(py) intermolecular interactions. *CrystEngComm* **2007,** *9* (1), 46-54.

159. Nangia, A.; Desiraju, G. R. Supramolecular synthons and pattern recognition. In *Design of Organic Solids*, 1998; pp 57-95.

160. Aakeroy, C. B.; Beatty, A. M.; Helfrich, B. A. "Total synthesis" supramolecular style: Design and hydrogen-bond-directed assembly of ternary supermolecules. *Angew. Chem. ,Int. Ed.* **2001,** *40* (17), 3240.

161. Allen, F. H.; Motherwell, W. D. S.; Raithby, P. R.; Shields, G. P.; Taylor, R. Systematic analysis of the probabilities of formation of bimolecular hydrogen-bonded ring motifs in organic crystal structures. *New J. Chem.* **1999,** *23* (1), 25-34.

162. Steiner, T. Competition of hydrogen-bond acceptors for the strong carboxyl donor. *Acta Crystallogr. , Sect. B* **2001,** *57*, 103-106.

163. Sreekanth, B. R.; Vishweshwar, P.; Vyas, K. Supramolecular synthon polymorphism in 2 : 1 co-crystal of 4-hydroxybenzoic acid and 2,3,5,6-tetramethylpyrazine. *Chem. Commun.* **2007,** (23), 2375-2377.

164. Fabian, L. Cambridge Structural Database Analysis of Molecular Complementarity in Cocrystals. *Cryst. Growth Des.* **2009,** *9* (3), 1436-1443.

165. Anderson, K. M.; Goeta, A. E.; Steed, J. W. Au center dot center dot center dot Au interactions: Z '> 1 behavior and structural analysis. *Inorg. Chem.* **2007,** *46* (16), 6444-6451.

166. Anderson, K. M.; Goeta, A. E.; Hancock, K. S. B.; Steed, J. W. Unusual variations in the incidence of Z ' > 1 in oxo-anion structures. *Chem. Commun.* **2006,** (20), 2138-2140.

167. Anderson, K. M.; Goeta, A. E.; Hancock, K. S. B.; Steed, J. W. Unusual variations in the incidence of Z ' > 1 in oxo-anion structures (pg 2138, 2006). *Chem. Commun.* **2006,** (25), 2722.

168. Aakeroy, C. B.; Beatty, A. M.; Helfrich, B. A.; Nieuwenhuyzen, M. Do polymorphic compounds make good cocrystallizing agents? A structural case study that demonstrates the importance of synthon flexibility. *Cryst. Growth Des.* **2003,** *3* (2), 159-165.

169. Zaworotko, M. Pharmaceutical co-crystals: do they represent a new path to improved medicines? *J. Pharm. Pharmacol.* **2006,** *58*, A91.

170. Chadwick, K.; Sadiq, G.; Davey, R. J.; Seaton, C. C.; Pritchard, R. G.; Parkin, A. Designing Acid Acid Co-Crystals-The Use of Hammett Substitution Constants. *Cryst. Growth Des.* **2009,** *9* (3), 1278-1279.

171. Hammett, L. P. The Effect of Structure upon the Reactions of Organic Compounds. Benzene Derivatives. *J. Am. Chem. Soc.* **1937,** *59* (1), 96-103.

172. Paulekuhn, G. S.; Dressman, J. B.; Saal, C. Trends in Active Pharmaceutical Ingredient Salt Selection based on Analysis of the Orange Book Database. *J. Med. Chem.* **2007,** *50* (26), 6665-6672.

173. Haleblian, J. K. Characterization of Habits and Crystalline Modification of Solids and Their Pharmaceutical Applications. *J. Pharm. Sci.* **1975,** *64* (8), 1269-1288.

174. U.S.Food and Drug Administration . Orange Book Database. 1981.
Ref Type: Catalog

175. Infantes, L.; Chisholm, J.; Motherwell, S. Extended motifs from water and chemical functional groups in organic molecular crystals. *CrystEngComm* **2003,** *5*, 480-486.

176. Copyright CCDC 2001-2007 *Mercury CSD 2.2 (Build RC5),* **2009**

177. Allen, F. H.; Raithby, P. R.; Shields, G. P.; Taylor, R. Probabilities of formation of bimolecular cyclic hydrogen- bonded motifs in organic crystal structures: a systematic database analysis. *Chem. Commun.* **1998,** (9), 1043-1044.

178. Desiraju, G. R. Supramolecular Synthons in Crystal Engineering - a New Organic Synthesis. *Angew. Chem. ,Int. Ed. Engl.* **1995,** *34* (21), 2311-2327.

179. Trask, A. V.; Motherwell, W. D. S.; Jones, W. Physical stability enhancement of theophylline via cocrystallization. *Int. J. Pharm.* **2006,** *320* (1-2), 114-123.

180. Almarsson, O.; Zaworotko, M. J. Crystal engineering of the composition of pharmaceutical phases. Do pharmaceutical co-crystals represent a new path to improved medicines? *Chem. Commun.* **2004,** *17*, 1889-1896.

181. Aakeroy, C. B.; Fasulo, M. E.; Desper, J. Cocrystal or salt: Does it really matter? *Molecular Pharmaceutics* **2007,** *4* (3), 317-322.

182. Haynes, D. A.; Jones, W.; Motherwell, W. D. S. Cocrystallisation of succinic and fumaric acids with lutidines: a systematic study. *CrystEngComm* **2006,** *8* (11), 830-840.

183. Stanton, M. K.; Bak, A. Physicochemical properties of pharmaceutical co-crystals: A case study of ten AMG 517 co-crystals. *Cryst. Growth Des.* **2008,** *8* (10), 3856-3862.

184. Johnson, S. L.; Rumon, K. A. Infrared Spectra of Solid 1:1 Pyridine-Benzoic Acid Complexes; the Nature of the Hydrogen Bond as a Function of the Acid-Base Levels in the Complex. *J. Phys. Chem.* **2008,** *69* (1), 74-86.

185. Childs, S. L.; Stahly, G. P.; Park, A. The salt-cocrystal continuum: The influence of crystal structure on ionization state. *Molecular Pharmaceutics* **2007,** *4* (3), 323-338.

186. Dale, S. H.; Elsegood, M. R. J.; Hemmings, M.; Wilkinson, A. L. The co-crystallisation of pyridine with benzenepolycarboxylic acids: The interplay of strong and weak hydrogen bonding motifs. *CrystEngComm* **2004,** *6*, 207-214.

187. Karki, S.; Friscic, T.; Jones, W.; Motherwell, W. D. S. Screening for pharmaceutical cocrystal hydrates via neat and liquid-assisted grinding. *Molecular Pharmaceutics* **2007,** *4* (3), 347-354.

188. Bruker AXS Inc. *SAINT+*, version 6.45; Madison, Wisconsin, USA, **2003**

189. Sheldrick, G. M. *SADABS (version 2.03) Program for Bruker Area Detector Absorption Correction,* University of Gottingen: Gottingen, Germany, **2001**

190. Sheldrick, G. M. *SHELXS97,* University of Göttingen: Göttingen, Germany, **1997**

191. Sheldrick, G. M. *SHELXL97*, University of Göttingen: Göttingen, Germany, **1997**

192. Cason, C.; Froehlich, T.; Kopp, N.; Parker, R. *Persistence of Vision Raytracer (POV-Ray),* version 3.6.1; Persistence of Vision Team: **2004**

193. Karamertzanis, P. G. *OptimalPaste, Cluster and Analysis,* **2005**

194. Hunter, B. *Rietica - A visual Rietveld program*, version 1.77; **2001**

195. Stone, A. J. Distributed multipole analysis: Stability for large basis sets. *J. Chem. Theory Comput.* **2005,** *1* (6), 1128-1132.

196. Allen, F. H.; Kennard, O.; Watson, D. G. Tables of bond lengths determined by x-ray and neutron diffraction. Part 1. Bond lengths in Organic compounds. *J. Chem. Soc. Perkin T. 2* **1987,** (12), S1-S19.

197. Saunders, V. R.; Dovesi, R.; Roetti, C.; Orlando, R.; Zicovich-Wilson, C. M.; Harrison, N. M.; Doll, K.; Civalleri, B.; Bush, I. J.; D'Arco, P.; Llunell, M. *Crystal06 User's Manual*, **2006**

198. Brown, H. C. *Determination of Organic Structures by Physical Methods;* Academic Press, New York: 1955.

199. Perrin, D. D. *Dissociation Constants of Organic Bases in Aqueous Solution;* Butteworths, London: 1965.

200. van Eijck, B. P.; Kroon, J. Structure predictions allowing more than one molecule in the asymmetric unit. *Acta Crystallogr. , Sect. B* **2000,** *56*, 535-542.

201. Chen, B.; Munson, E. J. Investigation of the mechanism of n-butane oxidation on vanadium phosphorus oxide catalysts: Evidence from isotopic labeling studies. *J. Am. Chem. Soc.* **2002,** *124* (8), 1638-1652.

202. Flitcroft, T.; Skinner, H. A.; Whiting, M. C. Heats of Hydrogenation .1. Dodeca-3-9 and Dodeca-5-7 Diynes. *Transactions of the Faraday Society* **1957,** *53* (6), 784-790.

203. Chatterjee, S.; Pedireddi, V. R.; Rao, C. N. R. Unexpected isomerization of maleic acid to fumaric acid on co-crystallization with 4,4 '-bipyridine. *Tetrahedron Lett.* **1998,** *39* (18), 2843-2846.

204. James, M. N. G.; Williams, G. J. B. A refinement of the crystal structure of maleic acid. *Acta Crystallogr. , Sect. B.* **1974,** *30* (5), 1249-1257.

205. Biradha, K.; Zaworotko, M. J. Supramolecular isomerism and polymorphism in dianion salts of pyromellitic acid. *Crystal Engineering* **1998,** 67-78.

206. Leiserowitz, L. Molecular Packing Modes. Carboxylic Acids. *Acta Crystallogr. , Sect. B.* **1976,** *32* (3), 775-802.

207. Stephenson, G. A. Anisotropic lattice contraction in pharmaceuticals: The influence of cryo-crystallography on calculated powder diffraction patterns. *J. Pharm. Sci.* **2006,** *95* (4), 821-827.

208. Haynes, D. A.; Pietersen, L. K. Hydrogen bonding networks in ammonium carboxylate salts: the crystal structures of phenylethylammonium fumarate-fumaric acid, phenylethylammonium succinate-succinic acid and anilinium fumarate-fumaric acid. *CrystEngComm* **2008,** *10* (5), 518-524.

209. MacDonald, J. C.; Dorrestein, P. C.; Pilley, M. M. Design of supramolecular layers via self-assembly of imidazole and carboxylic acids. *Cryst. Growth Des.* **2001,** *1* (1), 29-38.

210. Ballabh, A.; Trivedi, D. R.; Dastidar, P.; Suresh, E. Hydrogen bonded supramolecular network in organic salts: crystal structures of acid-base salts of dicarboxylic acids and amines. *CrystEngComm* **2002,** (24), 135-142.

211. Williams, D. E. Improved intermolecular force field for crystalline oxohydrocarbons including O-H...O hydrogen bonding. *J. Comput. Chem.* **2001,** *22* (1), 1-20.

212. Day, G. M.; Cooper, T. G.; Cruz Cabeza, A. J.; Hejczyk, K. E.; Ammon, H. L.; Boerrigter, S. X. M.; Tan, J.; Della Valle, R. G.; Venuti, E.; Jose, J.; Gadre, S. R.; Desiraju, G. R.; Thakur, T. S.; van Eijck, B. P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Neumann, M.; Leusen, F. J. J.; Kendrick, J.; Price, S. L.; Misquitta, A. J.; Karamertzanis, P. G.; Welch, G. W. A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; van de Streek, J.; Wolf, A.; Schweizer, B. Significant progress in predicting the crystal structures of small organic molecules - a report on the fourth blind test. *Acta Crystallogr. , Sect. B* **2009,** *65* (2), 107-125.

213. Fischer, A.; Vaughan, J.; Galloway, W. J. Structure + Reactivity in Pyridine Series .1. Acid Dissociation Constants of Pyridinium Ions. *J. Chem. Soc.* **1964,** (OCT), 3591-3596.

214. Albert, A.; Serjeant, E. P. *The Determination of Ionization Constants: A Laboratory Manual;* Cambridge University Press: 1984.

215. Cabeza, A. J. C.; Day, G. M.; Motherwell, W. D. S.; Jones, W. Prediction and observation of isostructurality induced by solvent incorporation in multicomponent crystals. *J. Am. Chem. Soc.* **2006,** *128* (45), 14466-14467.

216. Neumann, M. A.; Leusen, F. J. J.; Kendrick, J. A Major Advance in Crystal Structure Prediction. *Angew. Chem. ,Int. Ed.* **2008,** *47* (13), 2427-2430.

217. Price, S. L. Computed crystal energy landscapes for understanding and predicting organic crystal structures and polymorphism. *Accounts Chem. Res.* **2009,** *42* (1), 117-126.

218. Copyright CCDC 2001-2009 *Mercury CSD 2.3 (Build RC4),* **2009**

219. Issa, N.; Karamertzanis, P. G.; Welch, G. W. A.; Price, S. L. Can the Formation of Pharmaceutical Cocrystals Be Computationally Predicted? I. Comparison of Lattice Energies. *Cryst. Growth Des.* **2009,** *9* (1), 442-453.

220. Bleidner, W. E.; Harmon, J. B.; Hewes, W. E.; Lynes, T. E.; Hermann, E. C. Absorption Distribution and Excretion of Amantadine Hydrochloride. *Journal of Pharmacology and Experimental Therapeutics* **1965,** *150* (3), 484-490.

221. Forman, M. S.; Trojanowski, J. Q.; Lee, V. M. Y. Neurodegenerative diseases: a decade of discoveries paves the way for therapeutic breakthroughs. *Nature Medicine* **2004,** *10* (10), 1055-1063.

222. Witt, A.; Macdonald, N.; Kirkpatrick, P. Memantine hydrochloride. *Nat Rev Drug Discov* **2004,** *3* (2), 109-110.

223. Kemp, J. A.; McKernan, R. M. NMDA receptor pathways as drug targets. *Nature Neuroscience* **2002,** *5*, 1039-1042.

224. Parsons, C. G.; Danysz, W.; Quack, G. Memantine is a clinically well tolerated N-methyl-D-aspartate (NMDA) receptor antagonist - a review of preclinical data. *Neuropharmacology* **1999,** *38* (6), 735-767.

225. National Institute for Health and Clinical Excellence . Alzheimer's disease - donepezil, rivastigmine, galantamine and memantine (review) - Final appraisal document. 2006.
Ref Type: Catalog

226. Belangergariepy, F.; Brisse, F.; Harvey, P. D.; Butler, I. S.; Gilson, D. F. R. Structure of Adamantanamine Hydrochloride at 143-K. *Acta Crystallographica Section C-Crystal Structure Communications* **1987,** *43*, 756-759.

227. Harvey, P. D.; Gilson, D. F. R.; Butler, I. S. A Study of the Phase-Transition and Molecular-Motion in Adamantanamine Hydrochloride. *J. Phys. Chem.* **1987,** *91* (5), 1267-1270.

228. Merli, V.; Daverio, P.; Kovacsne-Mezei, A.; Aronhime, J. United States Patents Application Publication: Polymorphs of Memantine Hydrochloride. United States US 2006/0217573 A1, 2006.

229. Lou, W. J.; Hu, X. R.; Gu, J. M. Memantinium chloride 0.1-hydrate. *Acta Crystallographica Section E-Structure Reports Online* **2009,** *65*, O2191-U2342.

230. Hejczyk, K. E. Application of crystal structure prediction to salts and cocrystals. PhD thesis University of Cambridge, 2010.

231. Karamertzanis, P. G. *GenerateMoments*, **2005**

232. Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984,** *5* (2), 129-145.

233. Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990,** *11* (4), 431-439.

234. Bondi, A. Van der Waals Volumes + Radii. *J. Phys. Chem.* **1964,** *68* (3), 441-451.

235. Spek, A. L. *PLATON, A Multipurpose Crystallographic Tool*, Utrecht University: Utrecht, The Netherlands, **2003**

236. Vandersluis, P.; Spek, A. L. Bypass - An Effective Method for the Refinement of Crystal-Structures Containing Disordered Solvent Regions. *Acta Cryst. A* **1990,** *46*, 194-201.

237. Cruz-Cabeza, A. J.; Day, G. M.; Jones, W. Predicting Inclusion Behaviour and Framework Structures in Organic Crystals. *Chem. Eur. J* **2009,** *15* (47), 13033-13040.

238. Mohamed, S.; Barnett, S. A.; Tocher, D. A.; Shankland, K.; Leech, C. K.; Price, S. L. Discovery of three polymorphs of 7-fluoroisatin reveals challenges in using computational crystal structure prediction as a complement to experimental screening. *CrystEngComm* **2008,** *10*, 399-404.

239. Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko, A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooij, W. T. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. Crystal structure prediction of small organic molecules: a second blind test. *Acta Crystallogr. , Sect. B* **2002,** *58*, 647-661.

240. Jackson, R. A.; Mort, K. A. Computer modelling of complex molecular ionic materials. *Computational Materials Science* **2000,** *17* (2-4), 230-233.

241. Telfer, G. B.; Gale, J. D.; Roberts, K. J.; Jackson, R. A.; Wilde, P. J.; Meenan, P. A transferable interatomic potential for alkali chlorates and bromates. *Acta Crystallogr. , Sect. A.* **1997,** *53* (Pt4), 415-420.

242. Roberts, K. J.; Telfer, G. B.; Jackson, R. A.; Wilde, P. J.; Meenan, P. Determination of a Transferable Interatomic Potential For Alkali- Metal Perchlorates and Its Application to Morphological Modeling. *J. Chem. Soc. Faraday T.* **1995,** *91* (22), 4133-4138.

243. Hanke, C. G.; Price, S. L.; Lynden-Bell, R. M. Intermolecular potentials for simulations of liquid imidazolium salts. *Mol. Phys.* **2001,** *99* (10), 801-809.

244. Moult, J.; Fidelis, K.; Kryshtafovych, A.; Rost, B.; Hubbard, T.; Tramontano, A. Critical assessment of methods of protein structure prediction - Round VII. *Proteins-Structure Function and Bioinformatics* **2007,** *69*, 3-9.

245. Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. Predicting small-molecule solvation free energies: An informal blind test for computational chemistry. *J. Med. Chem.* **2008,** *51* (4), 769-779.

246. Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. A test of crystal structure prediction of small organic molecules. *Acta Crystallogr. , Sect. B* **2000,** *56*, 697-714.

247. Day, G. M.; Motherwell, W. D. S.; Ammon, H. L.; Boerrigter, S. X. M.; Della Valle, R. G.; Venuti, E.; Dzyabchenko, A.; Dunitz, J. D.; Schweizer, B.; van Eijck, B. P.; Erk, P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Leusen, F. J. J.; Liang, C.; Pantelides, C. C.; Karamertzanis, P. G.; Price, S. L.; Lewis, T. C.; Nowell, H.; Torrisi, A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; Verwer, P. A third blind test of crystal structure prediction. *Acta Crystallogr. , Sect. B* **2005,** *61* (5), 511-527.

248. Bock, H.; Van, T. T. H.; Schodel, H. Crystal structure of 1,8-naphthyridinium-(1)-tetraphenylborate - Flattening of a distorted molecular skeleton by protonation (short comm). *Monatshefte fur Chemie* **1996,** *127* (4), 391-396.

249. Shan, N.; Batchelor, E.; Jones, W. Quinolinium fumarate. *Acta Crystallogr. , Sect. E* **2003,** *59*, O397-O398.

250. Allen, F. H.; Kennard, O. *Chemical Design Automation News* **1993,** *8* (1), 31-37.